OLUME 105 PART B NUMBER 20 MARCH 1958

The Proceedings

OF

THE INSTITUTION OF ELECTRICAL ENGINEERS

FOUNDED 1871 : INCORPORATED BY ROYAL CHARTER 1921

PART B

RADIO AND ELECTRONIC ENGINEERING
(INCLUDING COMMUNICATION ENGINEERING)

SAVOY PLACE · LONDON W.C.2

Price Ten Shillings and Sixpence

THE INSTITUTION OF ELECTRICAL ENGINEERS

FOUNDED 1871 INCORPORATED BY ROYAL CHARTER 1921

PATRON: HER MAJESTY THE QUEEN

COUNCIL 1957–1958

President

T. E. GOLDUP, C.B.E.

Past-Presidents

SIR JAMES SWINBURNE, Bart., F.R.S.
W. H. ECCLES, D.Sc., F.R.S.
THE RT. HON. THE EARL OF MOUNT EDGCUMBE, T.D.
J. M. DONALDSON, M.C.
PROFESSOR E. W. MARCHANT, D.Sc.
H. T. YOUNG.
SIR GEORGE LEE, O.B.E., M.C.
SIR ARTHUR P. M. FLEMING, C.B.E., D.Eng., LL.D.
J. R. BEARD, C.B.E., M.Sc.
SIR NOEL ASHBRIDGE, B.Sc.(Eng.).
COLONEL SIR A. STANLEY ANGWIN, K.C.M.G., K.B.E., D.S.O., M.C., T.D., D.Sc.(Eng.).
SIR HARRY RAILING, D.Eng.
P. DUNSHEATH, C.B.E., M.A., D.Sc.(Eng.).
SIR VINCENT Z. DE FERRANTI, M.C.
T. G. N. HALDANE, M.A.
PROFESSOR E. B. MOULLIN, M.A., Sc.D.
SIR ARCHIBALD J. GILL, B.Sc.(Eng.).
SIR JOHN HACKING.
COLONEL B. H. LEESON, C.B.E., T.D.
SIR HAROLD BISHOP, C.B.E., B.Sc.(Eng.).
SIR JOSIAH ECCLES, C.B.E., D.Sc.
SIR GEORGE H. NELSON, Bart.
SIR GORDON RADLEY, K.C.B., C.B.E., Ph.D.(Eng.).

Vice-Presidents

S. E. GOODALL, M.Sc.(Eng.). WILLIS JACKSON, D.Sc., D.Phil., Dr.Sc.Tech., F.R.S. G. S. C. LUCAS, O.B.E.
SIR HAMISH D. MACLAREN, K.B.E., C.B., D.F.C., LL.D., B.Sc. C. T. MELLING, C.B.E., M.Sc.Tech.

Honorary Treasurer

THE RT. HON. THE VISCOUNT FALMOUTH.

Ordinary Members of Council

PROFESSOR H. E. M. BARLOW, Ph.D., B.Sc.(Eng.).
J. A. BROUGHALL, B.Sc.(Eng.).
C. M. COCK.
SIR JOHN DEAN, B.Sc.
B. DONKIN, B.A.
J. S. FORREST, D.Sc., M.A.
PROFESSOR J. GREIG, M.Sc., Ph.D.
E. M. HICKIN.
J. B. HIGHAM, Ph.D., B.Sc.
D. MCDONALD, B.Sc.
F. C. MCLEAN, C.B.E., M.Sc.
B. L. METCALF, B.Sc.(Eng.).
J. R. MORTLOCK, B.Sc.(Eng.).
H. H. MULLENS, B.Sc.
A. H. MUMFORD, O.B.E., B.Sc.(Eng.).
R. H. PHILLIPS, T.D.
D. P. SAYERS, B.Sc.
C. E. STRONG, O.B.E., B.A., B.A.I.
H. WATSON-JONES, M.Eng.
D. B. WELBOURN, M.A.
H. WEST, M.Sc.

Chairmen and Past-Chairmen of Sections

Measurement and Control:
H. S. PETCH, B.Sc.(Eng.).
*D. TAYLOR, M.Sc., Ph.D.

Radio and Telecommunication:
J. S. MCPETRIE, Ph.D., D.Sc.
*R. C. G. WILLIAMS, Ph.D., B.Sc.(Eng.).

Supply:
PROFESSOR M. G. SAY, Ph.D., M.Sc., F.R.S.E.
*P. J. RYLE, B.Sc.(Eng.).

Utilization:
J. VAUGHAN HARRIES.
*H. J. GIBSON, B.Sc.

Chairmen and Past-Chairmen of Local Centres

East Midland Centre:
J. D. PIERCE.
*H. L. HASLEGRAVE, M.A., Ph.D., M.Sc.(Eng.).

Mersey and North Wales Centre:
T. MAKIN.
*P. D'E STOWELL, B.Sc.(Eng.).

North Midland Centre:
A. J. COVENEY.
*W. K. FLEMING.

North-Eastern Centre:
T. W. WILCOX.
*J. CHRISTIE.

Southern Centre:
L. G. A. SIMS, D.Sc., Ph.D.
*H. ROBSON, B.Sc.

North-Western Centre:
F. R. PERRY, M.Sc.Tech.
*T. E. DANIEL, M.Eng.

Northern Ireland Centre:
C. M. STOUPE, B.Sc.
*DOUGLAS S. PARRY.

Western Centre:
J. F. WRIGHT.
*PROFESSOR G. H. RAWCLIFFE, M.A., D.Sc.

Scottish Centre:
E. O. TAYLOR, B.Sc.
*PROFESSOR F. M. BRUCE, M.Sc., Ph.D

South Midland Centre:
L. L. TOLLEY, B.Sc.(Eng.).
*C. J. O. GARRARD, M.Sc.

* Past Chairman.

MEASUREMENT AND CONTROL SECTION COMMITTEE 1957–1958

Chairman

H. S. PETCH, B.Sc.(Eng.).

Vice-Chairmen

J. K. WEBB, M.Sc.(Eng.), B.Sc.Tech.; PROFESSOR A. TUSTIN, M.Sc.

Past-Chairmen

DENIS TAYLOR, M.Sc., Ph.D.; W. BAMFORD, B.Sc.

Ordinary Members of Committee

J. BELL, M.C.
E. W. CONNON, B.Sc.(Eng.), M.Eng.
D. EDMUNDSON, B.Sc.(Eng.).
W. S. ELLIOTT, M.A.
C. G. GARTON.
PROFESSOR K. A. HAYES, B.Sc.(Eng.).
W. C. LISTER, B.Sc.
A. J. MADDOCK, D.Sc.
R. S. MEDLOCK, B.Sc.
G. A. W. SOWTER, Ph.D., B.Sc.(Eng.).
R. H. TIZARD, B.A.
M. V. WILKES, M.A., Ph.D., F.R.S.

And

The President (*ex officio*).
The Chairman of the Papers Committee.
PROFESSOR J. GREIG, M.Sc., Ph.D. (representing the Council).
W. GRAY (representing the North-Eastern Radio and Measurement Group).
E. ROSCOE, J.P. (representing the North-Western Measurement and Control Group).
P. R. HOWARD, Ph.D., B.Sc.(Eng.) (nominated by the National Physical Laboratory).
H. M. GALE, B.Sc.(Eng.) (representing the South Midland Radio and Measurement Group).

RADIO AND TELECOMMUNICATION SECTION COMMITTEE 1957–1958

Chairman

J. S. MCPETRIE, Ph.D., D.Sc.

Vice-Chairmen

G. MILLINGTON, M.A., B.Sc. M. J. L. PULLING, C.B.E., M.A.

Past-Chairmen

R. C. G. WILLIAMS, Ph.D., B.Sc.(Eng.). H. STANESBY.

Ordinary Members of Committee

A. J. BIGGS, Ph.D., B.Sc.
W. J. BRAY, M.Sc.(Eng.).
H. A. M. CLARK, B.Sc.(Eng.).
C. W. EARP, B.A.
V. J. FRANCIS, B.Sc.
R. J. HALSEY, C.M.G., B.Sc.(Eng.).
B. G. PRESSEY, M.Sc.(Eng.), Ph.D.
W. ROSS, M.A.
T. B. D. TERRONI, B.Sc.
D. R. TURNER, M.Eng.
F. WILLIAMS, B.Sc.
W. E. WILLSHAW, M.B.E., M.Sc.Tech.

And

The President (*ex officio*).
The Chairman of the Papers Committee.
PROF. H. E. M. BARLOW, Ph.D., B.Sc.(Eng.) (representing the Council).
E. H. COOKE-YARBOROUGH (Co-opted Member).
A. E. TWYCROSS (representing the North-Eastern Radio and Measurement Group).
N. C. ROLFE, B.Sc.(Eng.) (representing the Cambridge Radio and Telecommunication Group).
J. MOIR (representing the South Midland Radio and Measurement Group).
G. J. SCOLES, B.Sc.(Eng.) (representing the North-Western Radio and Telecommunication Group).

The following nominees of Government Departments:
Admiralty: CAPTAIN R. L. CLODE, R.N.
Air Ministry: GROUP CAPTAIN A. FODEN, B.Sc.Tech., R.A.F.
War Office: BRIGADIER J. D. HAIGH, O.B.E., M.A.

Secretary

W. K. BRASHER, C.B.E., M.A., M.I.E.E.

Deputy Secretary

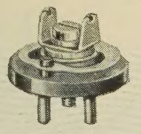
F. JERVIS SMITH, M.I.E.E.

Assistant Secretary

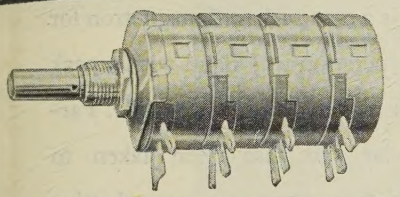
F. C. HARRIS.

Editor-in-Chief

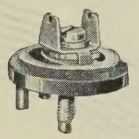
G. E. WILLIAMS, B.Sc.(Eng.), M.I.E.E.



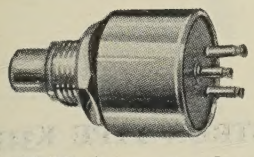
Preset Mk. 2 G type



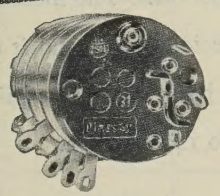
Type E four gang variation



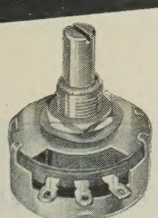
Preset Mk. 5 G type



MH2 Miniature panel and Spindle Sealed



Type E gang with Switch



Type EN to RCS 1000

Plessey

potentiometers

for special applications

Moulded track potentiometers by Plessey possess a high standard of stability within a wide range of operational temperatures and can be stored for extended periods without deterioration. They can be manufactured to conform, within strictly specified limits, to a designed pattern of values at various positions on the track.

COMMERCIAL TYPES

These are available as single units, dual concentrics or special gangs and are available with or without double pole switch, also in preset form. The standard resistance values are linear 100 ohms to 5 MΩ, log 5KΩ to 2 MΩ. The maximum working voltage is 500 V.

Standard Type 'E'. 2 watts at 70°C for linear laws.

Standard with double pole switch Type 'ES'. 2 watts at 70°C for linear laws.

Dual Type 'ED'. 2 watts at 70°C for each section.

Dual with double pole switch Type 'EDS'. 2 watts at 70°C for each section.

Preset Type 'EP'. 1 watt for 70°C for linear laws.

GOVERNMENT AND INSTRUMENTATION TYPES

Available as normal and preset types. In general are in accordance with the appropriate Ministry Specifications.

Panel and Spindle Sealed Type 'EH2' 1 watt at 100°C for linear laws. (Type Approval Certificate 906/4 [1230].) Maximum working voltage 500V.

Instrumentation Type 'EN' 2 watts at 70°C for linear laws. (All materials to R.C.S.1000.) Maximum working voltage 500V.

Miniature, thumb operated, with switch, Type 'F' ¼ watt at 70°C for linear laws. Maximum working voltage 250V.

Sub-Miniature Preset Type 'G'. Marks 2 & 5. ¼ watt at 70°C for linear laws. (Type Approval Certificates 974/2 and 1075/1 [1373].) Maximum working voltage 250V.

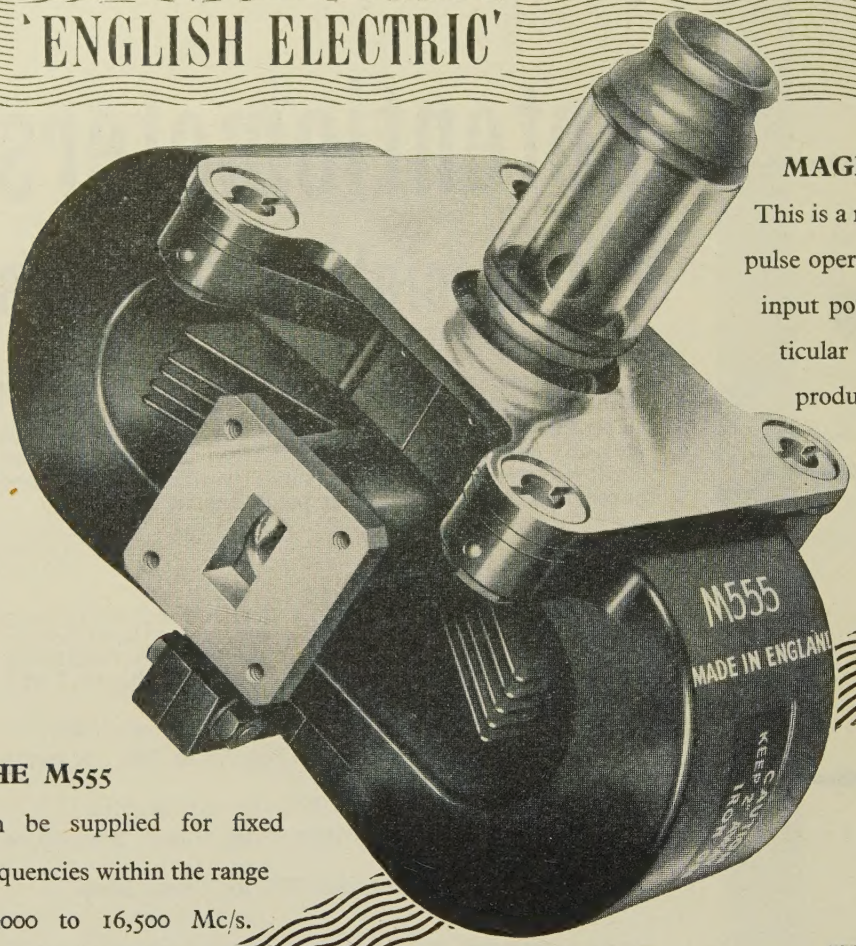
Miniature Panel and Spindle Sealed Type 'MH2' ¼ watt at 70°C. (Type Approval Certificate 1032/1 [1373/1].) Maximum working voltage 250V. A fully sealed version of this potentiometer is also available.

Rating at 100°C is normally one half of that at 70°C.

SWINDON COMPONENTS DIVISION
THE PLESSEY COMPANY LIMITED
KEMBREY STREET · SWINDON · WILTS · Tel: 5461
Overseas Sales Organisation: Plessey International Limited
Ilford · Essex · England · Telephone: Ilford 3040

THE VALVES FOR J BAND OPERATION

'ENGLISH ELECTRIC'



MAGNETRON TYPE M555

This is a new packaged magnetron for pulse operation in J Band with a peak input power rating of 240 kW. Particular care has been taken to produce a compact, rugged valve for air-borne applications.

THE M555

can be supplied for fixed frequencies within the range 14,000 to 16,500 Mc/s.

KLYSTRON TYPE K343

This is a low voltage reflex klystron for J Band operation with a minimum power output of 20 mW at 350 volts. The moulded base and flying leads specially commend it for high altitude operation. It has mechanical tuning covering the range 12,000 to 14,500 Mc/s.

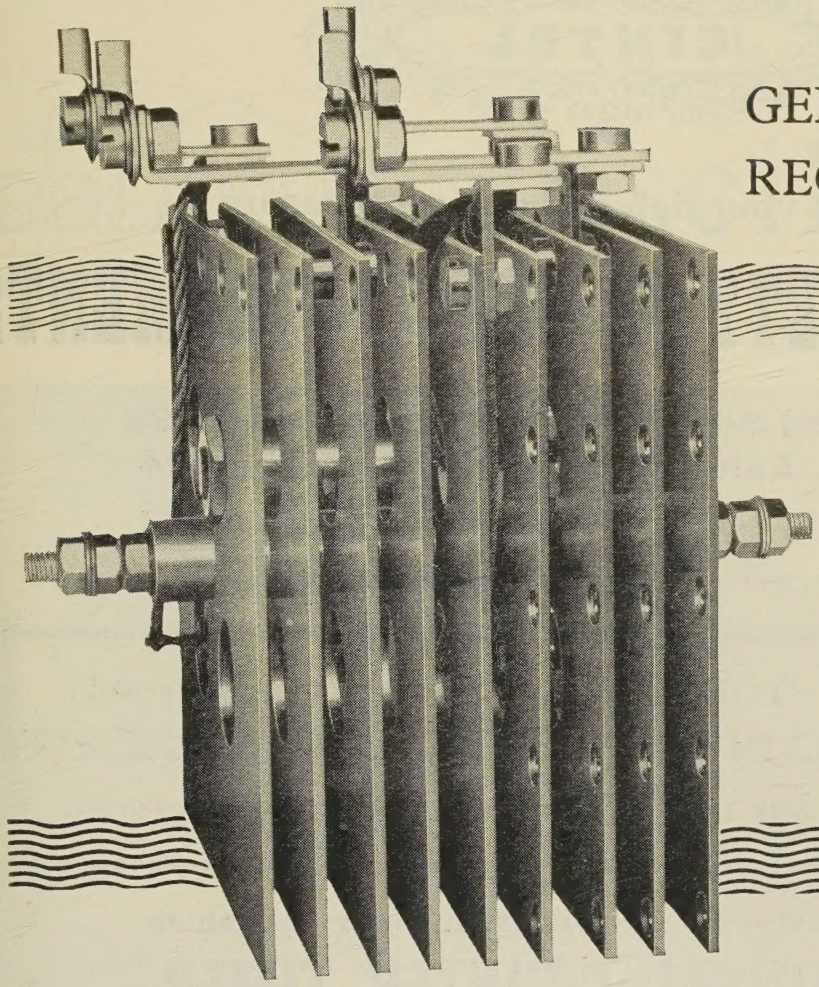
KLYSTRON TYPE K346

This is generally similar to type K343 with mechanical tuning from 14,500 to 17,000 Mc/s.

Both these klystrons, which may be used in conjunction with the M555 or in other J Band applications, have 30 to 80 Mc/s electronic tuning. The output connections are American type UG419/U feeding into No. 18 Waveguide.

ENGLISH ELECTRIC VALVE CO. LTD. EE *Chelmsford, England*
Telephone: Chelmsford 3491

AP/95



GERMANIUM RECTIFIERS

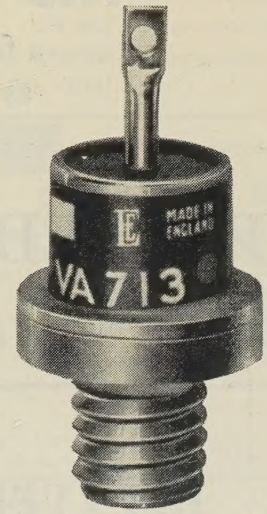
Due largely to advanced crystal perfection techniques and to the generous use of large area junctions, The English Electric Valve Co. Ltd. are able to offer Germanium Power Rectifiers and Stacks with high ratings at 55°C ambient temperature. Standard Stacks cater for all normal single and three-phase circuits.

Stacks are also designed on request to meet individual requirements.

Just a few of the possibilities of one E.E.V. Germanium Rectifier, the VA713, are listed in the table below.

Some Typical Stacked Assemblies

Stack No.	Circuit	R.M.S. Input Voltage per leg (Volts)	D.C. Output (Volts)	D.C. Output Current (Amperes)		No. of VA713 diodes	Approx. overall dimensions (inches)
				55°C.	35°C.		
13E1 13F1 13G1	Single phase, half wave	56 43 28	25 19 12	4·25 4·25 4·25	13 13 13	1 1 1	4 × 4⅝ × 2½
13E2 13F2 13G2	Single phase, full wave	28 21 14	25 19 12	8·5 8·5 8·5	26 26 26	2 2 2	4 × 4⅝ × 3⅜
13E3 13F3 13G3	Single phase, full wave bridge	56 43 28	50 38 24	8·5 8·5 8·5	26 26 26	4 4 4	4 × 4⅝ × 4⅝
13E4 13F4 13G4	Three phase, half wave	33 24 16	38 29 19	13 13 13	39 39 39	3 3 3	4 × 4⅝ × 4⅝
13E5 13F5 13G5	Three phase, full wave	33 24 16	75 56 38	13 13 13	39 39 39	6 6 6	4 × 4⅝ × 6¼
13E6 13F6 13G6	Six phase, half wave	28 21 14	38 29 19	26 26 26	78 78 78	6 6 6	4 × 4⅝ × 7¼



'ENGLISH ELECTRIC'

ENGLISH ELECTRIC VALVE CO. LTD. *Chelmsford, England*
Telephone: Chelmsford 3491

AP/99

let us demonstrate any of these superb SOLARTRON oscilloscopes

write or 'phone now for appointment

Have you seen the ***new*** Solartron Double-Beam Solarscope CD 711 being put through its paces?
Now, for the first time, you can choose a double-beam 'scope of Solartron quality, embodying all the latest design features and built to the finest standards of electronic and mechanical engineering.
Why not write or call us now for a demonstration of the Double-Beam Solarscope, or any of the eight other models listed below? Specialist instrument engineers are immediately available to assist you, whatever your problem or field of application.

	CD 711 DOUBLE BEAM
BANDWIDTH	Max. D.C.—7 Mc/s. at 100mV/cm.
SENSITIVITY	3mV/cm.—100 V/cm.
'Y' CALIBRATION	Cal. shift. Accuracy ±5%
'X' CALIBRATION	Cal. time scale. Accuracy ±5%
TIME BASE	Time scale 0.3μ Sec./cm.—3 Sec./cm.
'X' EXPANSION	Continuously variable ×10
SIZE & WEIGHT	16″ × 13″ × 27½″ deep. 110 lb.

	CD 715	AD 557	CD 518	CD 568	CD 513
BANDWIDTH	Max. D.C.—20 Kc/s.	Max. D.C.—1 Mc/s.	Max. D.C.—5 Mc/s.	Max. D.C.—5 Mc/s.	Max. D.C.—10 Mc/s.
SENSITIVITY	10 mV/cm.—10V/cm.	3 mV/cm.—100 V/cm.	0.25 V/cm.—5 V/cm.	0.25 V/cm.—5 V/cm.	1 mV/cm.—10 V/cm.
'Y' CALIBRATION	Special facilities	Cal. shift ±5%	Shift meter. Accuracy ±3%	Shift meter. Accuracy ±3%	Cal. sensitivity Accuracy ±10%
'X' CALIBRATION	Special facilities	Cal. time scale. Accuracy ±10%	'Cal Pips' and sine wave. Accuracy ±2%	Sine wave. Accuracy ±2%	Cal. time scale Accuracy ±10%
TIME BASE	Sweep time 10 Sec. to 0.1 Sec.	Time scale 1μ Sec./cm. —1 Sec./cm.	Sweep time 100m. Sec. —1μ Sec.	Sweep time 100m. Sec. —1μ Sec.	Time scale 0.1μ Sec./ cm.—1 Sec./cm.
'X' EXPANSION	×1, ×0.5, ×0.2, ×0.1, ×0.05	Continuously variable × 10			×0.5, ×1, ×2, ×5
SIZE & WEIGHT	14″ × 10″ × 20″ deep. 47 lb.	16½″ × 10″ × 22″ deep. 70 lb.	12″ × 9″ × 18″ deep. 40 lb.	12″ × 9″ × 18″ deep. 40 lb.	16¼″ × 10″ × 22″ deep. 70 lb.

	CD 523S	CD 814	CD 643
BANDWIDTH	Max. D.C.—10 Mc/s.	Constant 1 c/s.–8 Mc/s.	Constant D.C.— 15 Mc/s.
SENSITIVITY	1 mV/cm.—10 V/cm.	30 mV/cm.—30 V/cm.	100 mV/cm.–60 V/cm.
'Y' CALIBRATION	Cal. sensitivity Accuracy ±10%	Comparison method Accuracy ±5%	Cal. shift. Accuracy ±2%
'X' CALIBRATION	Cal. time scale Accuracy ±10%	'Cal Pips' Accuracy ±5%	'Cal Pips' and brightup Accuracy ±2%
TIME BASE	Time scale 0.1μ Sec./ cm.—1 Sec./cm.	Repetition Rate 6 c/s—185 Kc/s.	Time scale 0.1μ Sec/cm. —100m. Sec./cm.
'X' EXPANSION	× 0.5, × 1, × 2, × 5	Continuously variable × 10	Continuously variable × 100
SIZE & WEIGHT	16¼″ × 10″ × 23″ deep. 70 lb.	14½″ × 10½″ × 19½″ deep. 43 lb.	20″ × 14¾″ × 27¾″ deep. 140 lb.

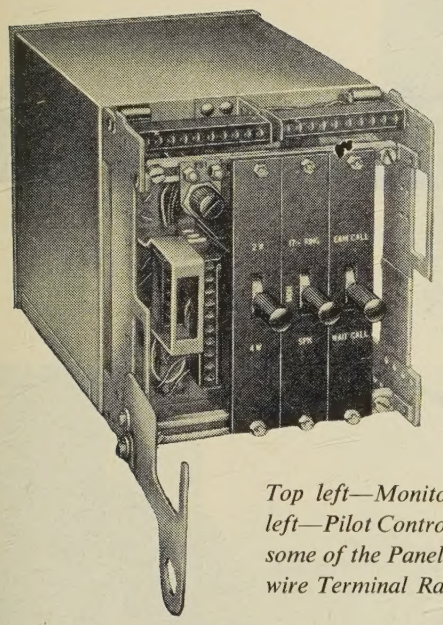
PRICES
Home & Export (F.O.B., packing extra)

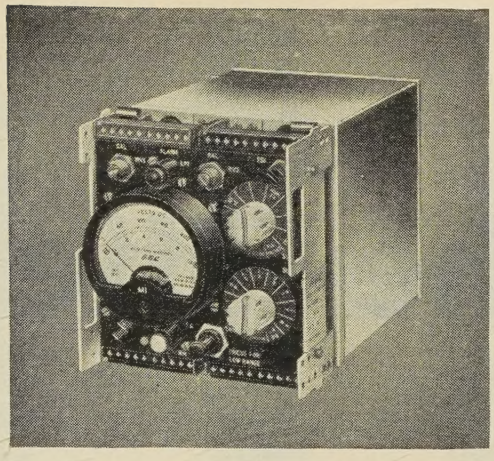
CD 711	£320
CD 715	£460
AD 557	£275
CD 518	£225
CD 568	£210
CD 513	£220
CD 523S	£265
CD 814	£115
CD 643	£490

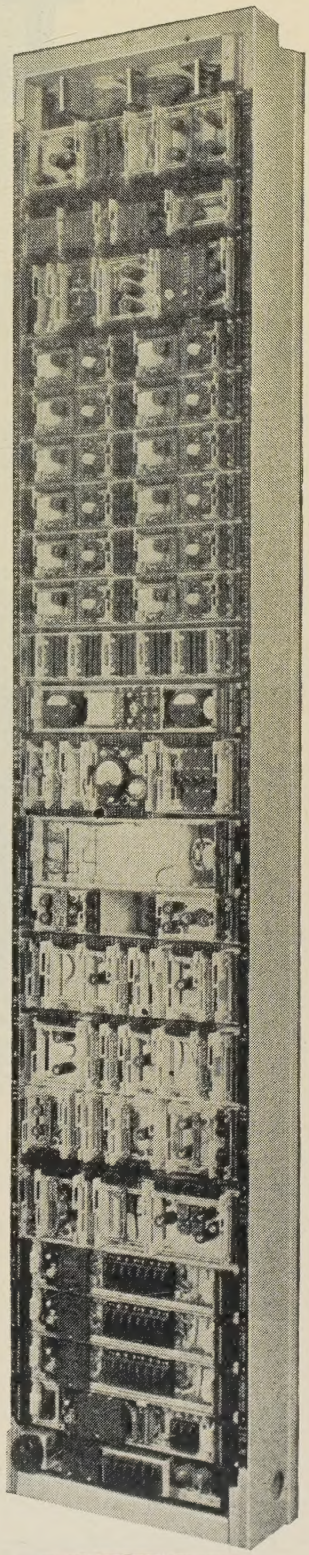
SOLARTRON

THE SOLARTRON ELECTRONIC GROUP LIMITED

THAMES DITTON • SURREY • Telephone: EMBerbrook 5522 • Cables: Solartron, Thames Ditton







Top left—Monitor, top right—Alarm, bottom left—Pilot Control and bottom right—Amplifier, some of the Panels used in the 12-Circuit Open-wire Terminal Rack shown on the right.

G.E.C. 3-circuit and 12-circuit carrier telephone equipment provides for the transmission of three and twelve high-quality speech circuits, respectively, over an open-wire route. Each speech circuit effectively transmits the frequency band 300 c/s to 3400 c/s. Signalling is effected at the out-of-band frequency of 3825 c/s on each circuit.

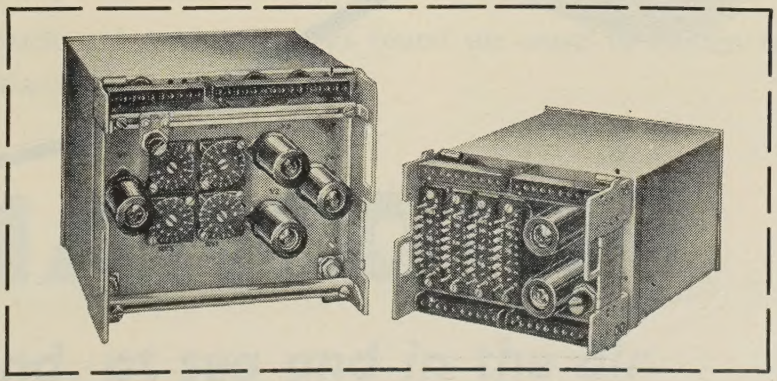
The signals transmitted to line in the 12-circuit system are within the frequency band 36 to 143 kc/s, and in the 3-circuit system within 3.16 to 31.1 kc/s in accordance with the recommendation of C.C.I.T.T. Thus a 3-circuit and a 12-circuit system can operate over the same open-wire pair. Several systems can operate along the same route.

When required, a broadcast programme circuit of approximately 10 kc/s bandwidth can be inserted at the intermediate frequency of 84 to 96 kc/s, in place of three speech circuits in either system. In addition, twenty-four voice-frequency telegraph channels can be provided over any speech circuit.

By suitably spacing repeaters along the route, a system may be operated over distances of many hundreds of miles, the exact distance and spacing of repeaters depending on local conditions.

The equipment itself is the latest G.E.C. Type 56. A complete 3-circuit terminal or a complete 12-circuit terminal is accommodated on one single-sided rack.

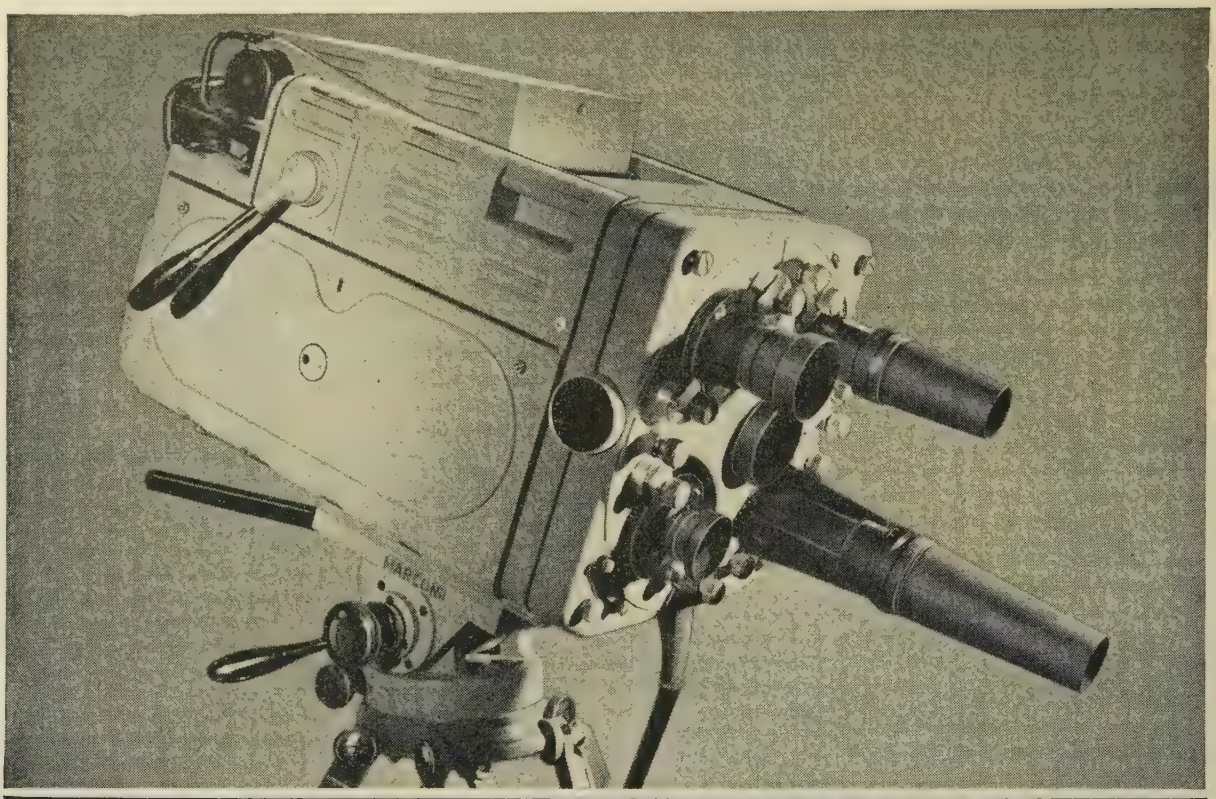
For further information please write for Standard Specifications
SPO1011
and
SPO1025



and TV link; short, medium and long haul. Automatic and Manual exchange.

TELEPHONE, RADIO & TELEVISION WORKS · COVENTRY · ENGLAND

G.E.C.



Marconi Camera Channels

IMAGE ORTHICON CAMERA Type BD808 (illustrated)

Features

- Uses either 3″ or 4½″ Image Orthicons.
- Designed for ease of servicing, excellent accessibility and plug-in sub-units.
- Four position turret will carry any combination from 2-inch to 40-inch lenses. 80-inch and zoom lenses may also be used.
- Viewfinder can be tilted up or down to give the most comfortable viewing position.
- Camera Control Unit may be used with 10″ picture tube and 3″ waveform tube, or with 14″ picture tube and 5″ waveform tube.
- Remote control of light intensity by variable graded filter.
- Optional remote control of focus and turret. Optional semi-automatic alignment circuit.
- Built-in turret for neutral density and colour filters.
- Full range of accessories available for both studio and outside broadcast roles.

BROADCAST VIDICON CAMERA Type BD364

The most recent addition to the Marconi range of Television Equipment.

Features

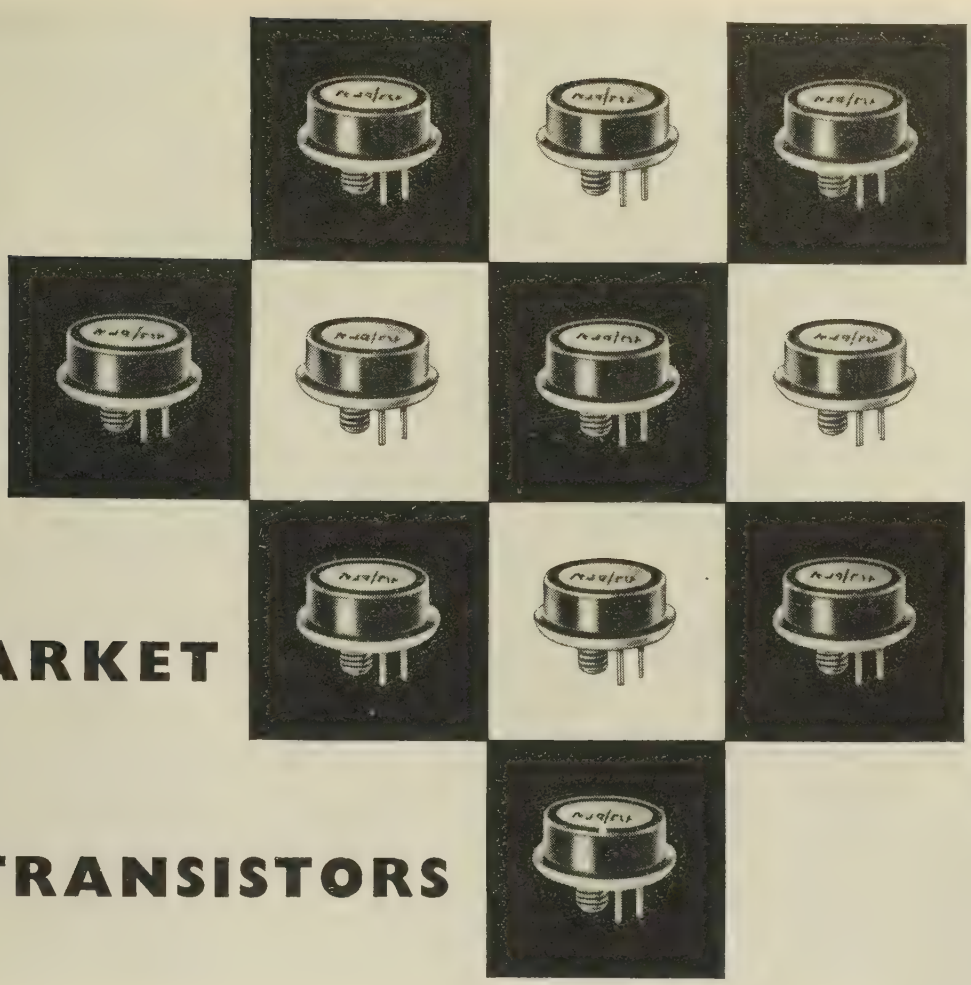
- Compact, easily operated by one man. The camera has integral viewfinder with 7″ tube and 2¾″ waveform monitor and includes all operational controls.
- Channel consists of Camera and Power Supply only—but optional Remote C.C.U. and Monitor position available.
- Use of close-tolerance double-triodes in all valve circuits except one and printed wiring assemblies ensures great reliability.
- Rapid semi-automatic beam alignment, built-in aperture correction and gamma correction circuits are provided. Designed to make the best use of any of the present Vidicon tubes and with ample flexibility to deal with foreseeable developments.
- 4-position turret with positive location takes wide range of fixed and zoom lenses.

MARCONI

COMPLETE SOUND BROADCASTING AND TELEVISION SYSTEMS

MARCONI'S WIRELESS TELEGRAPH COMPANY LIMITED, CHELMSFORD, ESSEX, ENGLAND

LB9



NEWMARKET TRANSISTORS

APPLICATIONS

Stabilized Power Packs
Relay devices
Relay operating devices
DC Motor control
DC Generator control
DC Converters
Sinusoidal Oscillators
Relaxation Oscillators

V30/10LP GOLTOP

POWER TRANSISTOR 25/-

MAXIMUM RATINGS

Collector Voltage 30V
Collector Current 3.0A
Junction Temperature 75°C
Collector Power Dissipation 10W
on 50 sq. in. of 16 s.w.g. Aluminium
Thermal Resistance 2°C/Watt
(*Junction to stud*)
Derating 200mW/°C rise above 25°C

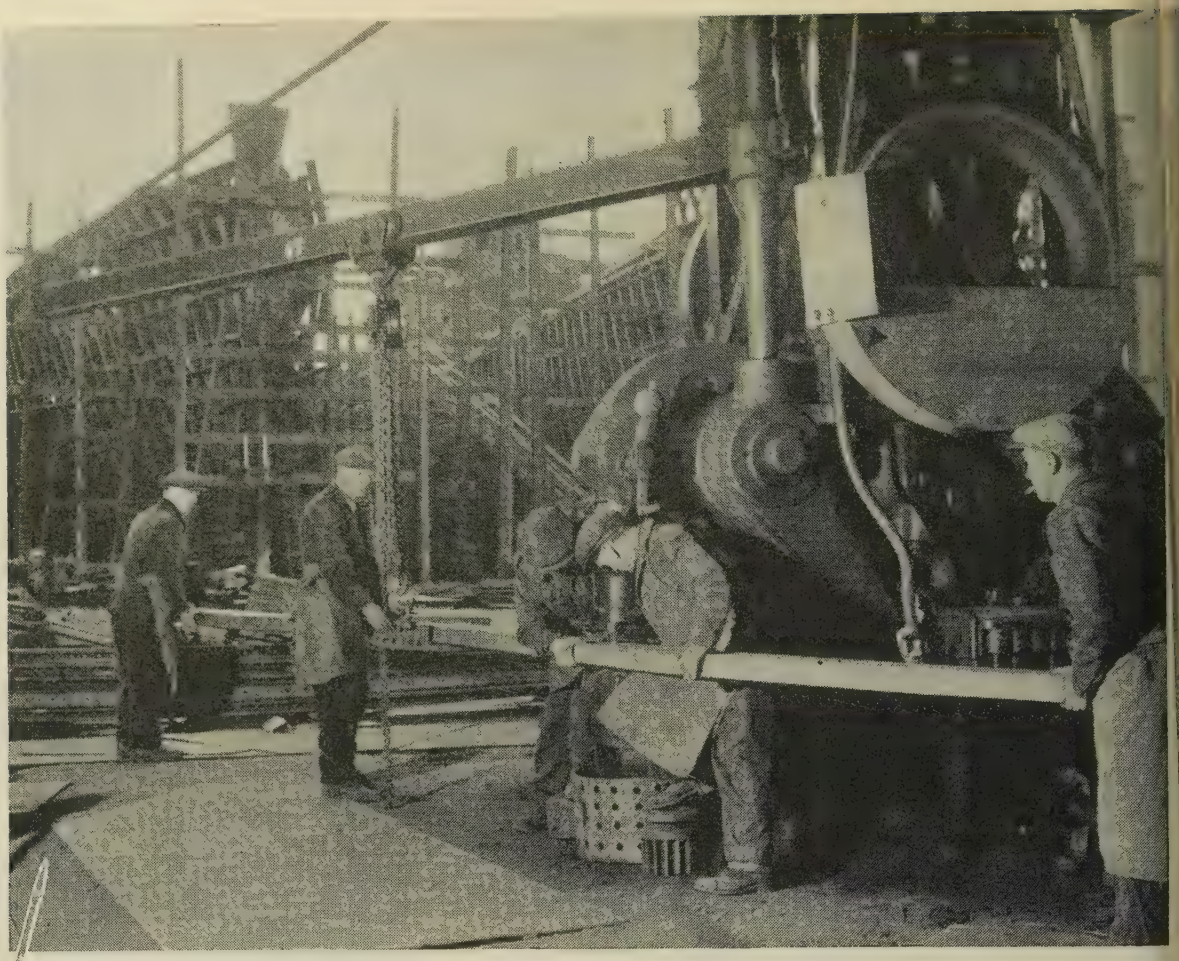
Visit us at STAND 142 at the R.E.C.M.F. and STAND 503 at the I.E.A. Exhibitions.

CHARACTERISTICS

Collector cut-off current (—1.5V) 30μA
Current Gain (—200mA) Minimum 10
Current Gain (—1.5A) 6

NEWMARKET TRANSISTOR COMPANY LIMITED

EXNING ROAD • NEWMARKET • SUFFOLK • *Telephone: Newmarket 3381/4*



Building a ship. Rivet holes are being punched in the frames to match the holes which will be punched in the shell.

C.M.A. Cables at work

In the busy shipyards of the Clyde, in factories all over the country, at home or abroad, wherever you go C.M.A. cables are at work. In the air, underground or underwater, in ships, coalmines and powerhouses, C.M.A. cables are transmitting power, unobtrusively . . . reliably . . . year in year out. For over 100 years members of the Cable Makers Association have been concerned in all major advances in cable making. Together, member firms of C.M.A. spend over one million pounds a year on research and development. At every stage from raw material to finished cable, technical knowledge is pooled and research co-ordinated to avoid wasted effort. This highly-organized co-operative research activity has contributed largely to the world-wide prestige that C.M.A. cables enjoy. It has put Britain at the head of the world's cable exporters. Technical information and advice is available from any member.

MEMBERS OF THE C.M.A.

British Insulated Callender's Cables Ltd. Connollys (Blackley) Ltd. Enfield Cables Ltd. W. T. Glover & Co. Ltd. Greengate & Irwell Rubber Co. Ltd. W. T. Henley's Telegraph Works Co. Ltd. Johnson & Phillips Ltd. The Liverpool Electric Cable Co. Ltd. Metropolitan Electric Cable & Construction Co. Ltd. Pirelli-General Cable Works Ltd. (The General Electric Co. Ltd). St. Helens Cable & Rubber Co. Ltd. Siemens Edison Swan Ltd. Standard Telephones & Cables Ltd. The Telegraph Construction & Maintenance Co. Ltd.

Insist on a cable with the C·M·A label

CABLE MAKERS ASSOCIATION, 52-54, High Holborn, London, W.C.1. Telephone: HOLborn 7633

CMA 16

a new valve for 470 Mc/s equipment

–unique double tetrode range extended

QQVO2-6

Here is a new six watt double tetrode for low cost 470Mc/s mobile equipment. This compact valve features a frame grid and the same unique twin construction as other Mullard double tetrodes — a construction which provides high efficiencies, high power gain and heater economy. Other features of the QQVO2-6 include built-in neutralising capacitors which enable circuitry to be simplified, and low inter-electrode capacitances which allow wide tuning ranges to be achieved and which contribute to high efficiency.
Write to the address below for full details of the QQVO2-6 and other Mullard double tetrodes.

ASE B9A
ATHODE ... ... Indirectly heated
EATER (Centre tapped)
ries 12.6V, 0.3A
rallel 6.3V, 0.6A

APACITANCES

a–g′ (each section)	less than 0.16pF
′–all (each section)	6.4 pF
–all (each section)	1.6 pF
ut (two sections in push-pull)	0.95pF
(two sections in push-pull)	3.8 pF

nternally neutralised for push-pull operation.

CHARACTERISTICS
(each section) measured at Ia = 25mA, Va = Vg″ = 150V

gm...	10.5mA/V
μg′–g″	31

TYPICAL OPERATING CONDITIONS

	Telegraphy or F.M.	Telephony — A.M.	
f	470	470	Mc/s
Va	180	180	V
Vg″	180	180	V
Ia	2 x 27.5	2 x 20	mA
pa	2 x 2.1	2 x 1.5	W
Pout	5.8	4.2	W
Pload	4.5	3.4	W

THE MULLARD DOUBLE TETRODE RANGE
. . . the most comprehensive and efficient in the world

	Typical F.M. Power Output
QQVO2–6	5.8 watts
QQVO3–10/6360 (CV2798) ...	11 watts
QQVO3–20A/6252 (CV2799) ...	25 watts
QQVO6–40A/5894 (CV2797) ...	56 watts

Mullard
COMMUNICATIONS AND INDUSTRIAL VALVE DEPARTMENT

MULLARD LIMITED · MULLARD HOUSE
TORRINGTON PLACE · W.C.1 · Tel: LANgham 6633
MVT343

UNDOUBTEDLY . . .

" For inductance or capacitance measurements the measuring bridge is energised by the output from a valve oscillator, the out-of-balance voltage from the bridge being applied to the built-in selective amplifier-detector and moving-coil indicator, while for resistance measurements the indicator is used directly as a centre-zero galvanometer, the necessary d.c. for the bridge being derived from the power pack supplying the oscillator and detector used for inductance and capacitance measurements."

One of the problems of the age is how to keep up with the younger generation. One can support the shame of being unable to help with homework, but the prospect of having to admit ignorance of electronic equipment is intolerable. For instance, what do you know about the Marconi Universal Bridge Type TF 868A? Could you impress our young friend with an authoritative description of the instrument's ingenious mechanical design which provides single-dial measurement of L, C, and R? You owe it to yourself to be really up-to-date about Marconi Instruments—after all, they are important tools of your trade. Start with the TF 868A. Our leaflet K112 contains full details, and we'll gladly send you a copy.

UNIVERSAL BRIDGE
Type TF 868A

Measures inductance or capacitance at 1 or 10 kc/s, resistance at d.c. Measurement Ranges: 1 μH to 100 henrys, 1 μμF to 100 μF, 0.1 ohm to 10 MΩ. Q Range: 0.1 to 10 at kc/s, 1 to 100 at 10 kc/s. Tan δ Range: 0.001 to 0.1 at 1 kc/s, 0.01 to 1.0 at 10 kc/s.

MARCONI INSTRUMENTS

AM & FM SIGNAL GENERATORS · AUDIO & VIDEO OSCILLATORS · FREQUENCY METERS · VOLTMETERS POWER METERS · DISTORTION METERS · FIELD STRENGTH METERS · TRANSMISSION MONITORS DEVIATION METERS · OSCILLOSCOPES, SPECTRUM & RESPONSE ANALYSERS · Q METERS & BRIDGES

MARCONI INSTRUMENTS LTD · ST. ALBANS · HERTFORDSHIRE · TELEPHONE: ST. ALBANS 56161

London and the South: Marconi House, Strand, London, W.C.2. Tel: COVent Garden 1234
Midlands: Marconi House, 24 The Parade, Leamington Spa. Tel: 1408 *North:* 30 Albion Street, Kingston-upon-Hull. Tel: Hull Central 16347
WORLD-WIDE REPRESENTATION

TC 112

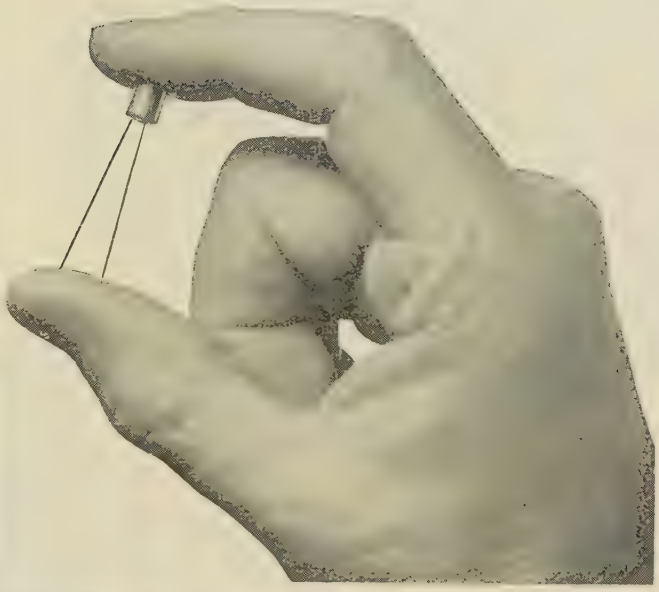
FERRANTI SILICON RECTIFIERS

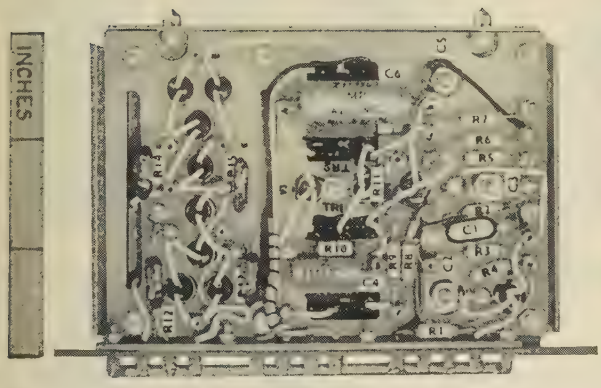
are used in

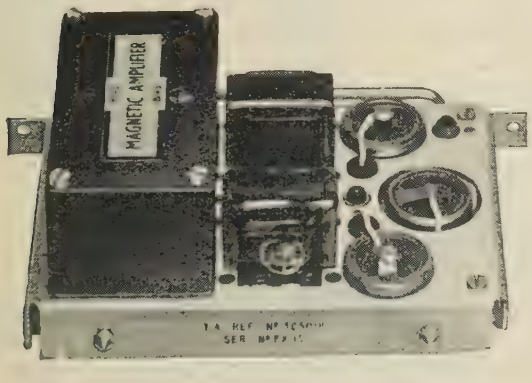
ELLIOTT

MAGNETIC AMPLIFIERS

Visit our Stand No. M9 at the A.S.E.E. Exhibition, March 25th to 29th







A wide range of applications

Ferranti Silicon Rectifiers are used in Magnetic Amplifiers made by Elliott Bros. (London) Ltd. They were chosen not only for their efficient operation, but also for their complete reliability, robust construction, small size and lightness in weight. Ferranti Silicon Rectifiers have many applications in the aircraft, electronic, electrical and general engineering industries including aero engine controls, aircraft power supplies, radar systems, guided missiles, computers, indicating and recording instruments, process control and telephone equipment.

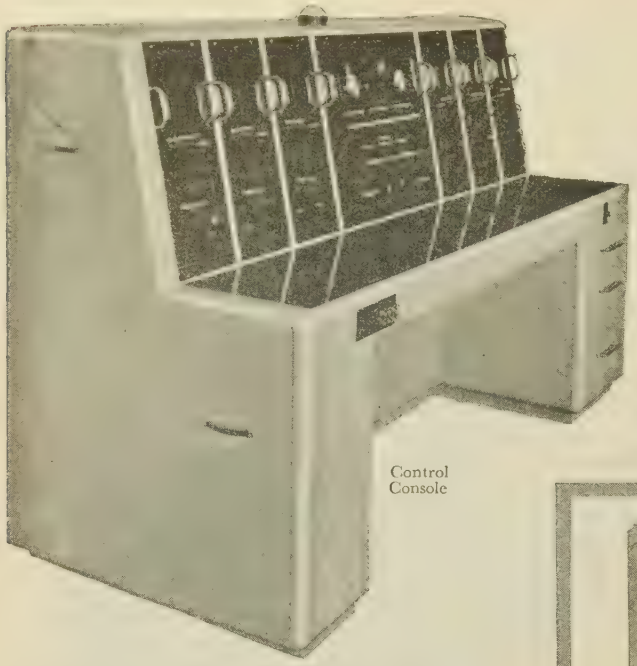
- ★ LOW REVERSE CURRENT
- ★ HIGH FORWARD SLOPE
- ★ HIGH TEMPERATURE OPERATION
- ★ SMALL PHYSICAL DIMENSIONS
- ★ HIGH MECHANICAL STRENGTH
- ★ APPROVED TO GOVERNMENT SPECIFICATIONS

FERRANTI LTD · GEM MILL · CHADDERTON · OLDHAM · LANCS

London Office: KERN HOUSE, 36 KINGSWAY, W.C.2

FERRANTI

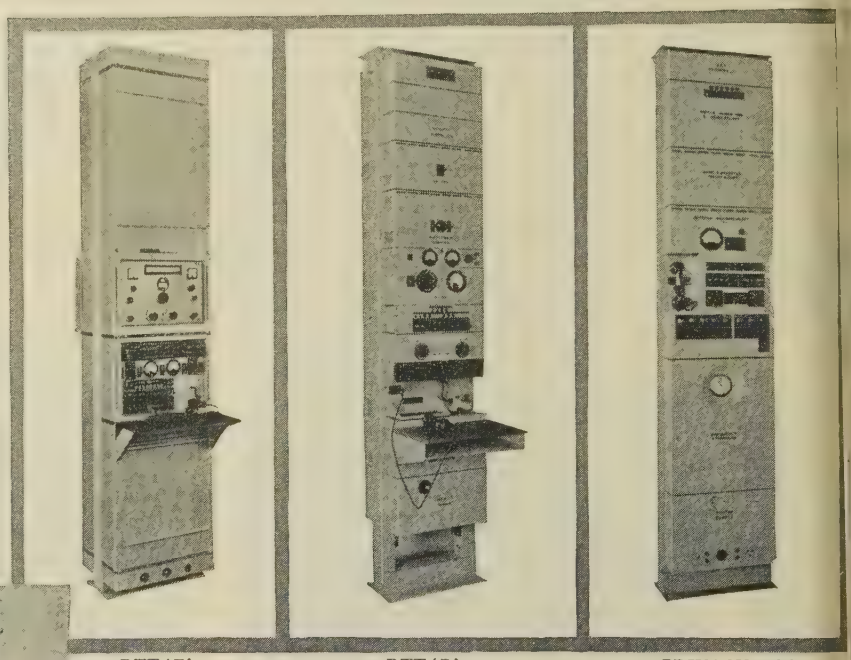
FE 177



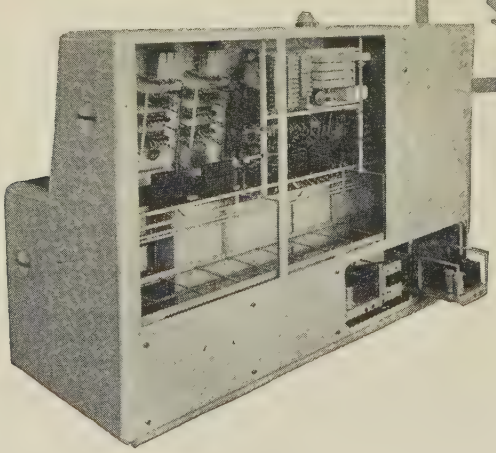
Control Console

simple or complex

* Electronically controlled 5 band privacy equipment.
* Channel shifting equipment for I.S.B. transmission.
* Centralised control, local or remote.
* Full monitoring and supervision facilities.



RTT 'F' RTT 'B' PRIVACY '5E'



Siemens Ediswan radio telephone terminal equipment provides efficient connection of a land telephone system to a radio transmitting and receiving station.

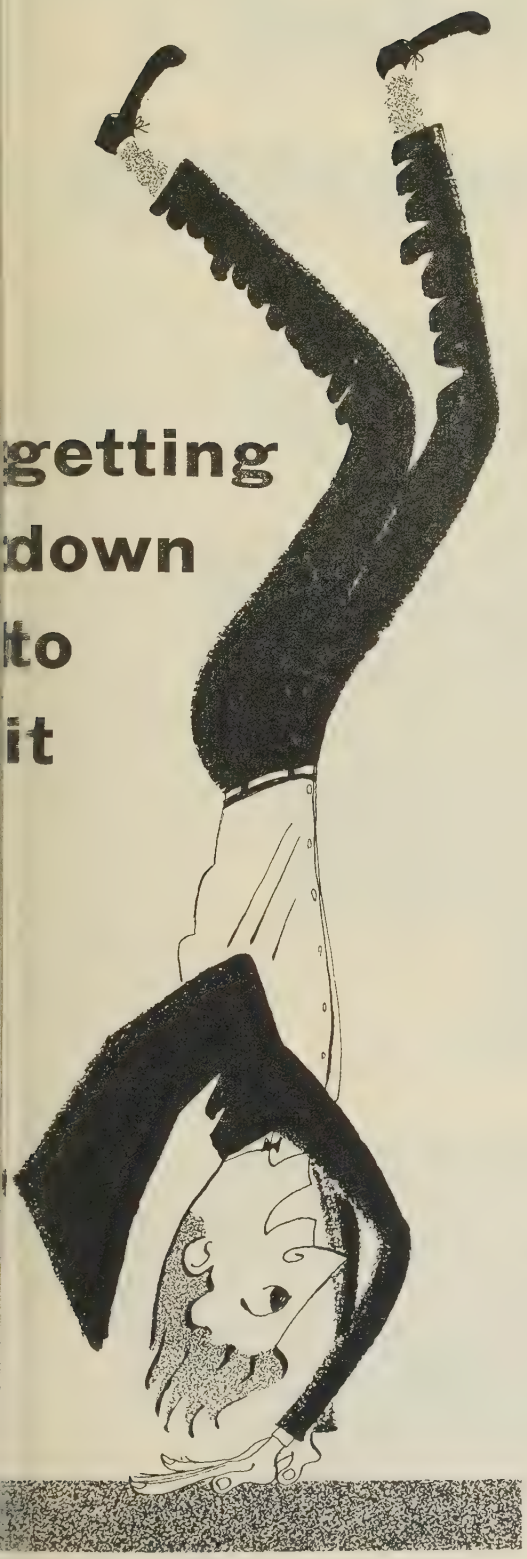
Privacy equipment of various types is obtainable, and allows a high degree of protection against unauthorised interception of speech on a radio telephone circuit.

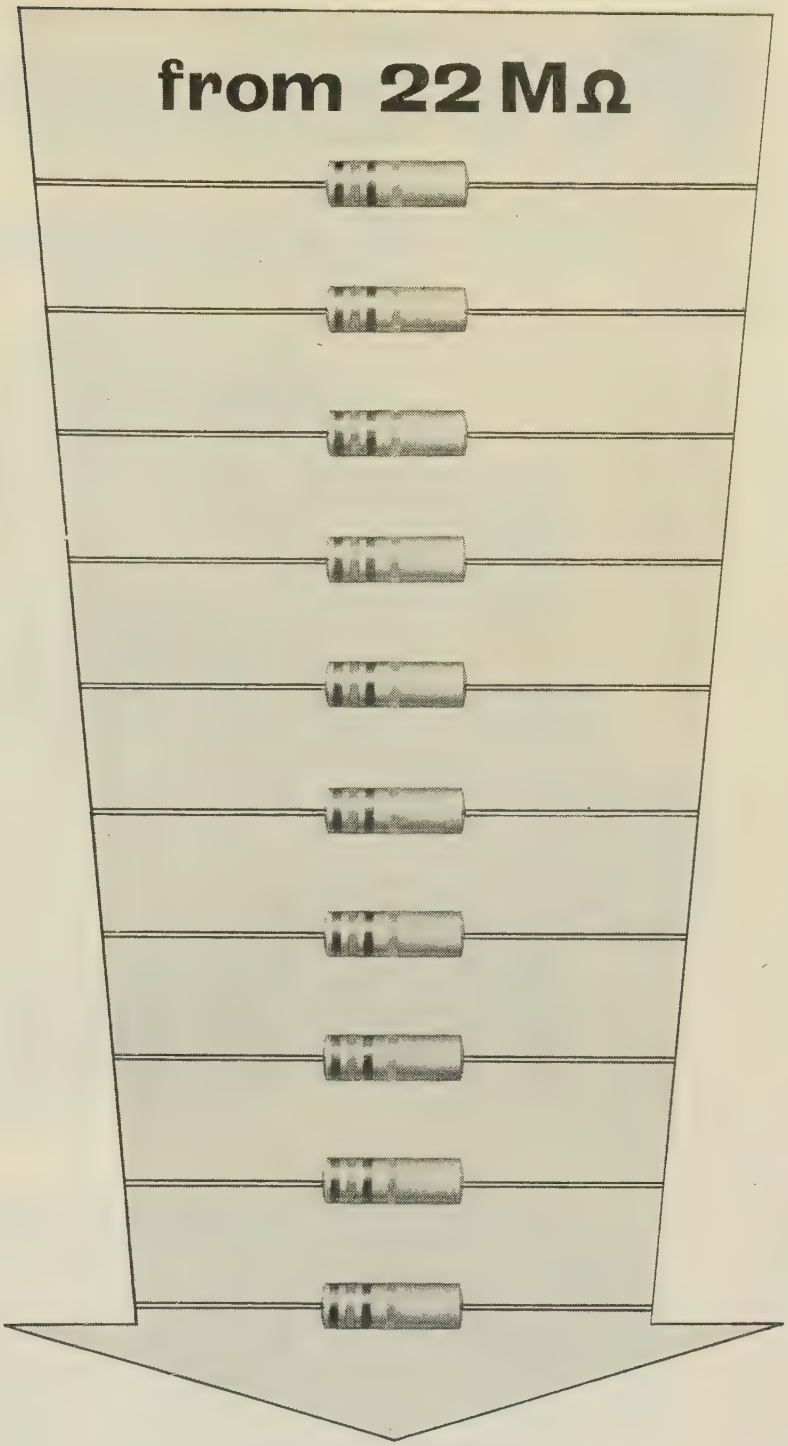
There is a wide range of equipment available. Full information will gladly be sent on request.

extending → **the frontiers of telecommunications**

SIEMENS EDISWAN

SIEMENS EDISON SWAN LTD. *An A.E.I. Company*
Telecommunications Transmission Division, Woolwich, London, S.E.18. Telephone: Woolwich 2020
Telegrams and Cables: Sieswan Souphone London.
This equipment can also be obtained from Marconi's Wireless Telegraph Company Ltd., Chelmsford, Essex
T.AB209





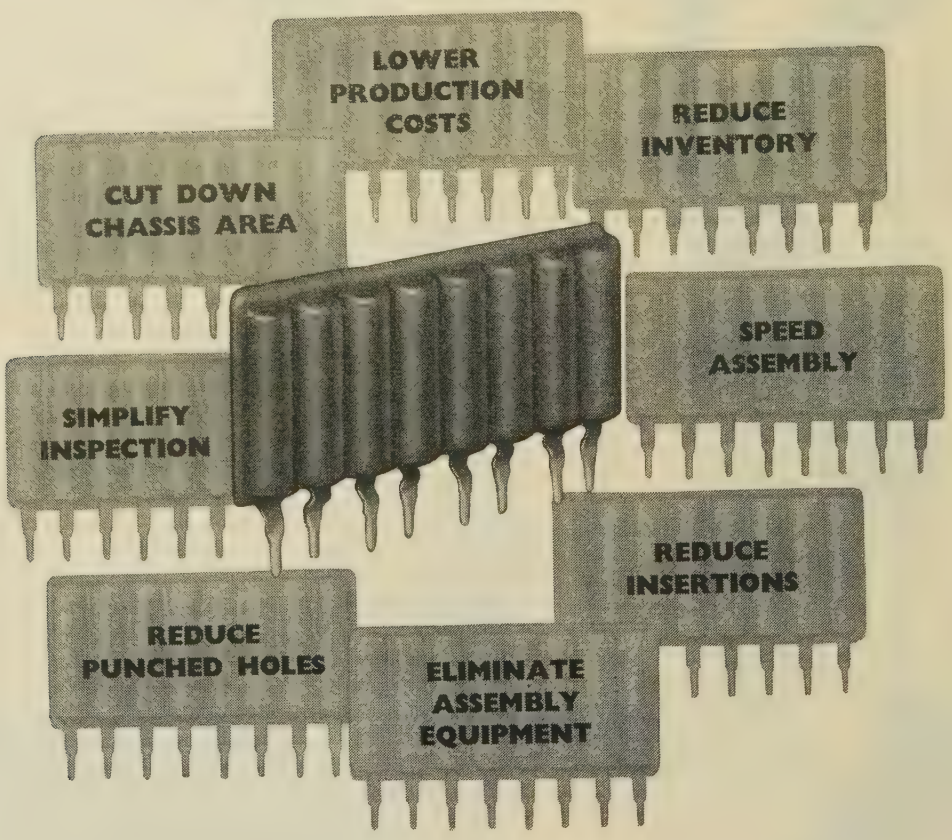
We are pleased to announce that our range of BTA (1W.) and BTS ($\frac{1}{2}$W). resistors has been extended down to 10Ω. We can now supply these resistors from stock in the preferred values from 10Ω to 22MΩ.

Please write for full details.

DUBILIER

DUBILIER CONDENSER CO. (1925) LTD., DUCON WORKS, VICTORIA ROAD, NORTH ACTON, LONDON W.3.
'hone: ACOrn 2241
Grams: Hivoltcon Wesphone London.
DN 204

Pre-assembled Components



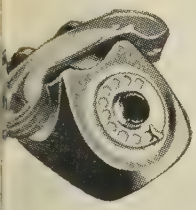
The Erie pre-assembled component system, known as "Pac", is a thin vertical module containing standard resistors and capacitors, of proven quality, mounted to a printed wiring board in a stable mechanical assembly, tailored to the requirements of the individual customer.

Besides offering considerable savings in itself, "Pac" paves the way for fully three-dimensional assembly, thus enabling the designer to take the fullest possible advantage of the potential savings in space attendant upon the introduction of shorter, wide-angle, tubes for television.

ERIE* *Resistor Ltd*

★ Registered Trade Mark

Carlisle Road, The Hyde, London, N.W.9., England. Tel: COLindale 8011. *Factories: London and Gt. Yarmouth, England; Toronto, Canada; Erie, Pa., and Holly Springs, Miss., U.S.A.*



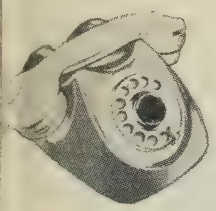
THE CENTENARY NEOPHONE

A quality instrument at an economical price, especially suitable for use overseas.

circuit telephone

This entirely new instrument, a product of the designers of the world-famous "Neophone" which established new standards in telephone performance and appearance in the nineteen-thirties, incorporates the following attractive features —

Improved performance, using latest design of components.
Lower first cost, achieved by the most modern manufacturing methods.
Lower maintenance cost.
Suitably finished for tropical use. Sealed case keeps out dust and insects.
Reduced size and weight—handset weight halved—only 7 ounces.



21 COLOUR COMBINATIONS

7 different handset shades and a choice of 3 case colours. The case can be changed WITHOUT DISTURBING THE DIAL SWITCH.

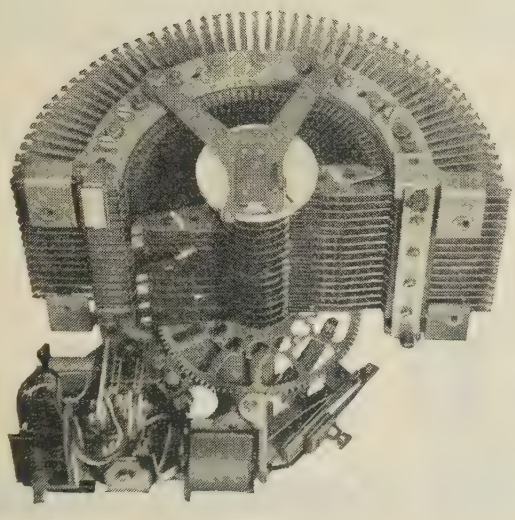
The centenary NEOPHONE

THE NEW INSTRUMENT WITH NEW FEATURES

Complete technical details are available on request.

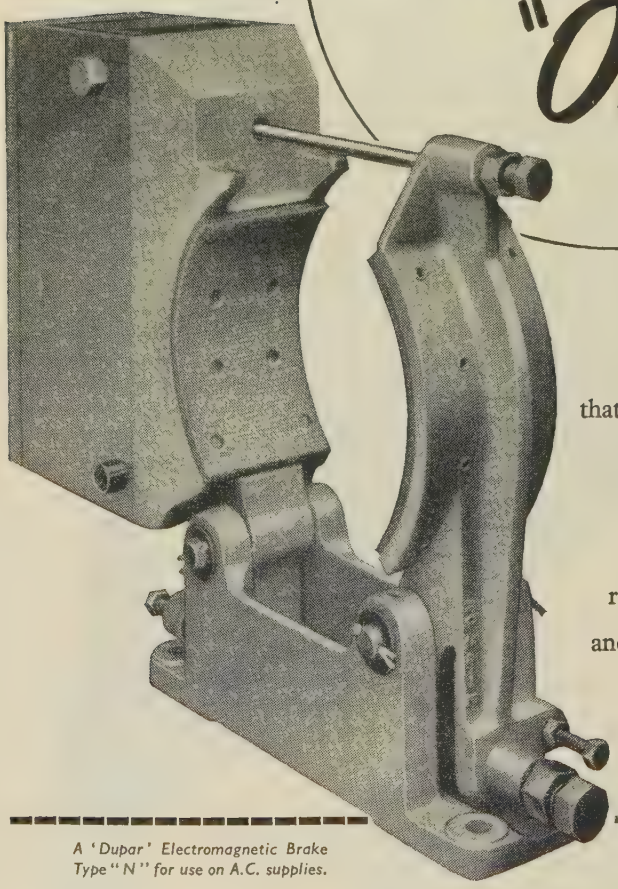
The equipment racks include every refinement for ease of maintenance—automatic call trace facilities, ribbon bank multiples, etc.—and are totally enclosed to minimize the effects of dust. The racks illustrated are 10 ft. 6 ins. in height.

An alternative unit construction is available which often proves more convenient for locations with restricted height or means of entry. The units illustrated are 8 ft. 9 ins. in height and this equipment has found much favour for private installations.



The basic switching element in our No. 17 System is the High Speed Motor Uniselector with a freedom from faults proved to be at least ten times greater than that of a bimotional mechanism. Such features as twin contact wipers, a vibrationless drive, and the elimination of wiper cords and complex mechanical actions, combine to ensure this very high degree of reliability.

"Over-run" means "Overtime"



A 'Dupar' Electromagnetic Brake
Type "N" for use on A.C. supplies.

Let any machine just roll to a stop and only overtime can keep up the output, for every unnecessary second lost in idle time is at the expense of production. Controlled braking is the answer—braking that becomes *immediately* and *automatically* operative on switching-off, and requires no complicated mechanical linkage. Such conditions are fully met by Dewhurst Electromagnetic Brakes. They are available in a wide range of types and sizes for all classes of service where safe and rapid braking is an essential or a desirable requirement.

We shall be pleased to advise on the correct choice of equipment on receipt of detailed information.

DUPAR

DEWHURST

ELECTROMAGNETIC BRAKES

for dependability

DEWHURST & PARTNER LTD · INVERNESS WORKS · HOUNSLOW · MIDDLESEX

Telephone: HOUnslow 0083 (8 lines) *Telegrams:* Dewhurst, Hounslow

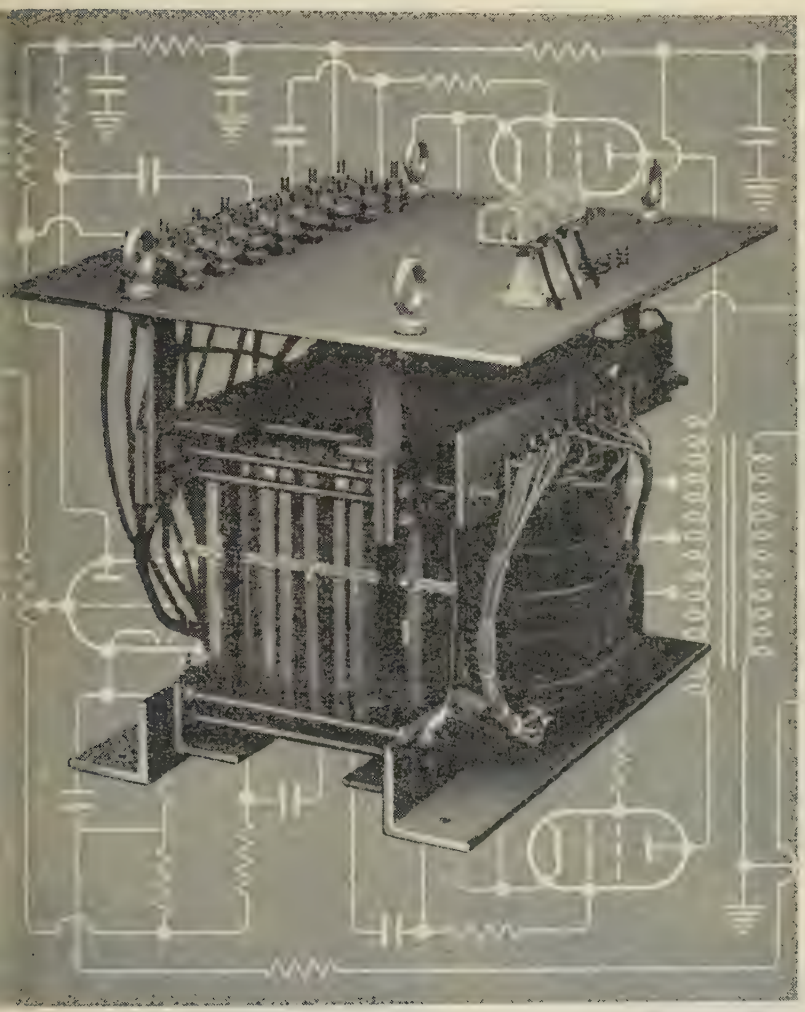
Branches: BIRMINGHAM · GLASGOW · LEEDS · MANCHESTER · NEWCASTLE · NOTTINGHAM

Revolutionary or Traditional

Some Massicore transformers break new ground, for Savage of Devizes are pioneers in this field and specialise in solving unusual problems.

Others play less spectacular roles, but *all* Massicore transformers have the same essential ingredient-conscientious craftsmanship. Massicore transformers are built to last a lifetime.

Equally important is the individual attention given to all enquiries and orders regardless of size; and we make a point of keeping our delivery promises.



The special transformer on the left required a core weighing no less than 216 lbs. In fact, the heaviest of the "standard" range of sizes of Grain Oriented Core Materials proved inadequate, and special tools and designs for larger sizes of our own had to be made.

The only really *special* feature about the simple choke illustrated above is the Savage craftsmanship which characterises all our products. It is choke type IG4—10 hys measured at 10v 50 c/s with 250 m/A. D.C. D.C. Resistance 50 ohms.

Your requirements may call for instruments very different from these examples. Please take advantage of our experience, knowledge and constructional skill in the production of all types of transformers.

Corner for Contented Customers

"... First of all, we would like to compliment you on your transformer which gives a very good performance in our instruments ..."

T. C. LONDON

"On a recent visit to the United States, I came across a considerable amount of enthusiasm for your products. ."

A. T. LONDON

MASSICORE SAVAGE

SAVAGE TRANSFORMERS LIMITED Devizes, Wiltshire Tel: Devizes 932

TP/59

Improved cable clamps on the new range of our Miniature Hermetically Sealed Co-axial Plugs and Sockets will withstand a pull of up to 35 lbs., and are suitable for use with Uniradio 70 cable.

They have been developed and Type Approved to Inter-Service Requirement under R.C.S. 322 and are freely interchangeable with the types which they supersede.

The new types are hermetically sealed and 100% production tested. Sealing caps are available for uncoupled units.

MAGNETIC DEVICES LIMITED

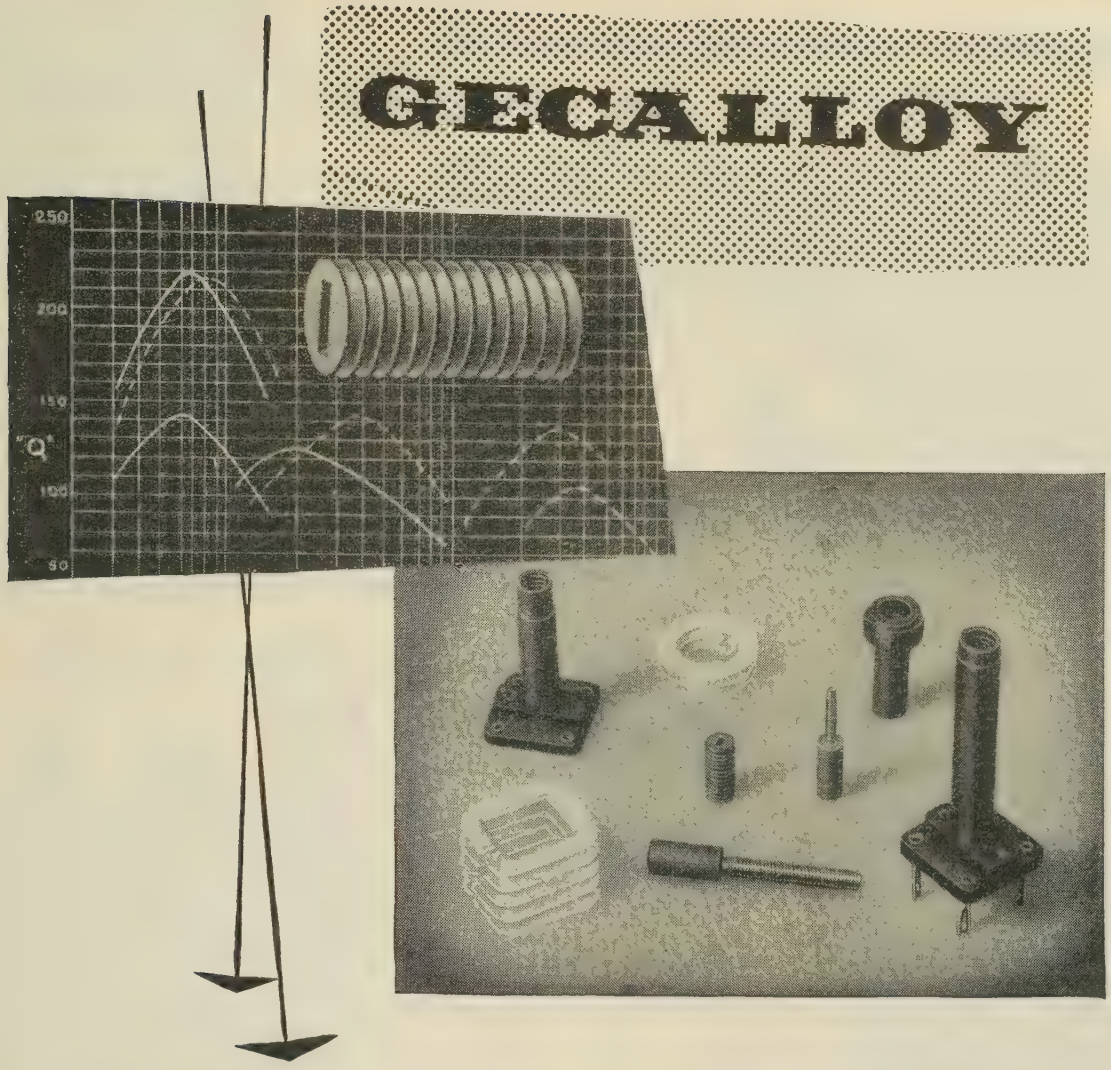
MAGNETIC DEVICES LIMITED

A. I. D. & A. R. B.—Approved

Exning Road, Newmarket, Suffolk

Phone: Newmarket 3181/2/3. Grams: Magnetic Newmarket.

MD 23



screw cores

Gecalloy screw cores for Radio and Television provide an inexpensive and simple method of inductance adjustment.

Moulding in three types of insulated powdered iron produces cores with the highest resistivity, ideal for H.F. and V.H.F. working.

Available in three grades: M.E. for frequencies from 100 Kc/s to 50 Mc/s
M.F. for frequencies from 5 Mc/s to 200 Mc/s
M.A. for frequencies from Mc/s to 300 Mc/s.

Write for leaflets giving full technical details.

ALFORD ELECTRICAL INSTRUMENTS LIMITED

(COMPONENTS GROUP)

IMES MILL · HEYWOOD · LANCASHIRE · Telephone Heywood 6868

London Sales Office: Telephone Temple Bar 4669

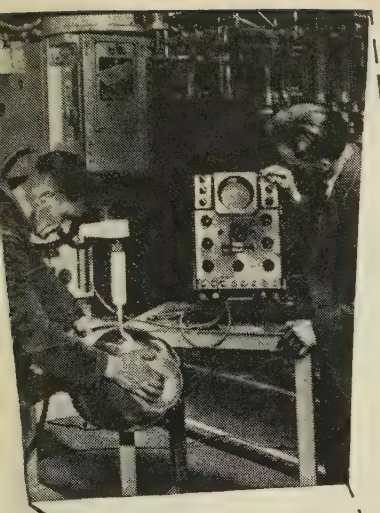
SUBSIDIARY OF THE GENERAL ELECTRIC CO. LTD. OF ENGLAND

THE PROCEEDINGS OF THE
INSTITUTION OF ELECTRICAL ENGINEERS

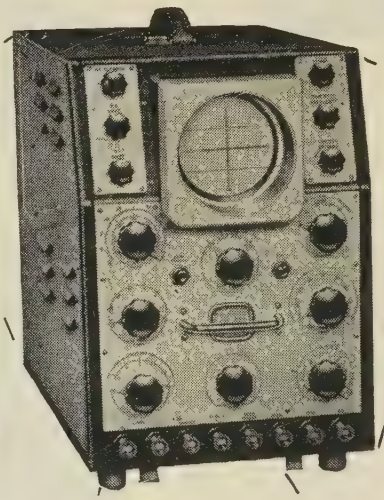
TEN YEAR INDEX
1942—1951

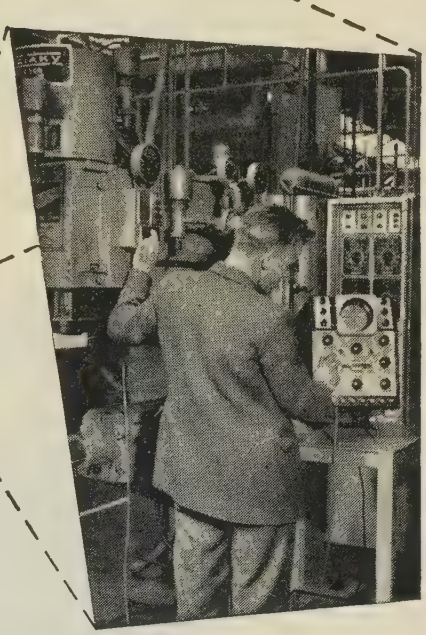
A TEN-YEAR INDEX to the *Journal of The Institution of Electrical Engineers* for the years 1942–48 and the *Proceedings* 1949–51 (vol. 89–98) can be obtained on application to the Secretary.

The published price is £1 5s. 0d. (post free), but any member of The Institution may have a copy at the reduced price of £1 (post free).



INVALUABLE IN EVERY FIELD OF INDUSTRY





COSSOR OSCILLOGRAPHS

Trial-and-error methods of fault evaluation are outmoded, inaccurate and expensive. With a Cossor Oscillograph, the modern Engineer is armed with an instrument that will display, measure and time the action of his electrical or mechanical machinery with precision, speed and certainty. It is worthwhile, therefore, to investigate the possibilities of these versatile instruments in *your* Industry by writing to:

COSSOR INSTRUMENTS LIMITED

The Instrument Company of the Cossor Group

COSSOR HOUSE · HIGHBURY GROVE · LONDON, N.5

Telephone: CANonbury 1234 (33 lines) Telegrams: Cossor, Norphone, London Cables: Cossor, London

(CL104)

We're skating on thick ice

When we say that we can execute your orders to the highest standards and in the best materials, with uniform quality throughout, and deliver them in bulk on or before time, we are not just making empty claims without the means or intention of carrying them out.
Behind these words we have an organisation designed specifically to do just this and it's doing *just this* at the moment for the many well known concerns that we have on our books. It can do it for you too.

As supplied, in a special heat resisting phenolic material, to The Zenith Carburetter Company Limited.

CERTIFIED ZINC ALLOY ZAMCA DIE CASTERS

Metropolitan Plastics Limited

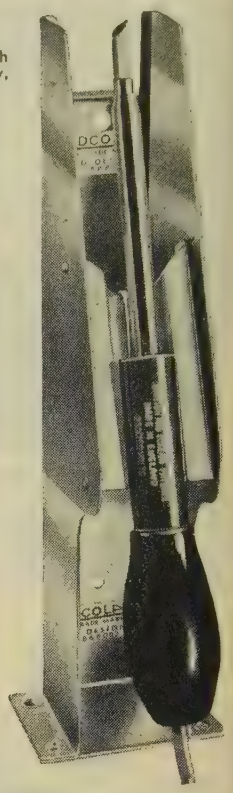
Specialists in thermo-setting plastics

Glenville Grove, Deptford, London SE8
Telephone: TIDeway 1172

ILLUSTRATED

$\frac{3}{16}$ **DETACHABLE BIT MODEL** (Factory bench line assembly, etc.) (List No. 64.)

PROTECTIVE SHIELD (List No. 68).



SOUND SOLDER JOINTING

Uniformity of shape, brightness of colour confirms permanent joints in any **SOUND UNIT.**

Sound Units are made with **SOUND JOINTS**

Sound Joints are only made by using

ADCOLA
(Regd. Trade Mark)
PATENTED

SOLDERING INSTRUMENTS and EQUIPMENT

Catalogues

Head Office, Sales and Service

ADCOLA PRODUCTS LTD.
GAUDEN ROAD
CLAPHAM HIGH STREET
LONDON, S.W.4

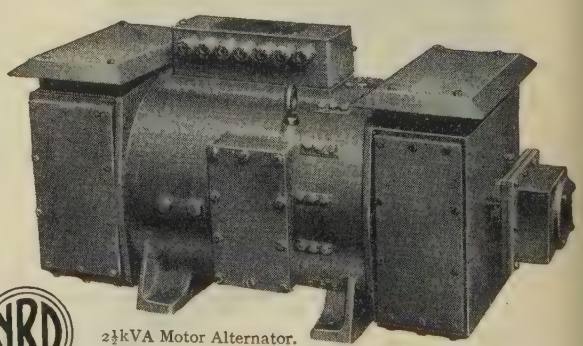
Telephones:
MACaulay 4272 & 3101

NEWTON-DERBY
ELECTRICAL EQUIPMENT

High Frequency Alternators

(Send for Publication No. 1003/2)

Also makers of Rotary Transformers and Converters, Wind and Engine-Driven Aircraft Generators, High Tension D.C. Generators, and Automatic Carbon Pile Voltage Regulators.

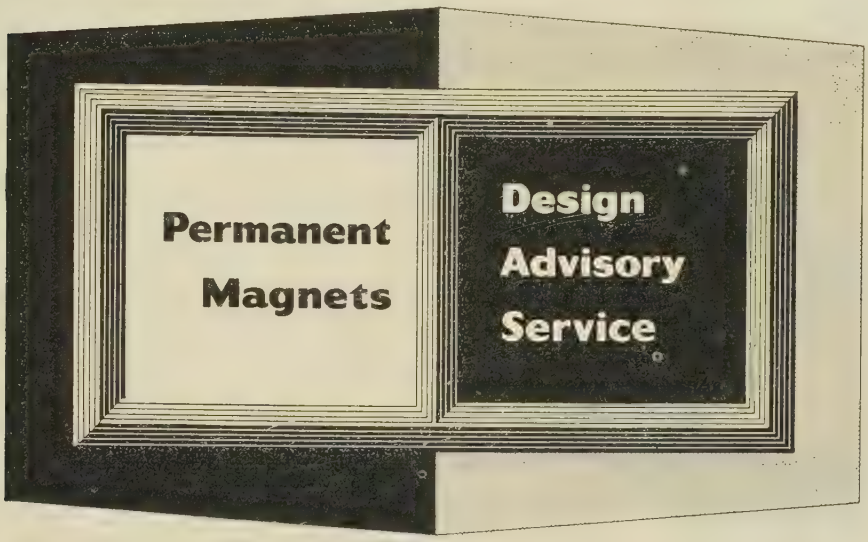


NBD

2½kVA Motor Alternator. Drip proof to 45°. Motor 220 volts D.C. Output 120 volts. 3 phase. 333 cycles per second. Motor includes an automatic constant speed governor. Weight 450 lb.

NEWTON BROTHERS
(DERBY) LTD

HEAD OFFICE & WORKS: ALFRETON ROAD, DERBY
TELEPHONE: DERBY 47676 (4 lines) TELEGRAMS: DYNAMO, DERBY
LONDON OFFICE: IMPERIAL BUILDINGS, 56 KINGSWAY W.C.2



Facilities for technical co-operation in the utilisation of magnet materials

An increasing number of equipment designers are taking advantage of the Mullard Design Advisory Service on permanent magnets. Through this service they can obtain specialist assistance on the selection, design and application of permanent magnets—assistance that will help them ensure optimum magnet performance in the finished equipment. The service is available to all users of permanent magnets, so if you have a design problem just write or telephone the address below: a team of specialist engineers will be at your disposal.

To help with the many day-to-day problems that designers encounter, the Mullard Design Advisory Service is preparing a series of advertisements covering a wide range of subjects from Modern Magnetic Theory to the application of Magnets for Nuclear Research. A new advertisement will be published monthly and each will be readily identified by the symbol shown above. If you would like to collate reprints of the series for reference, just send us a postcard and we will see that you receive them regularly as they become available.

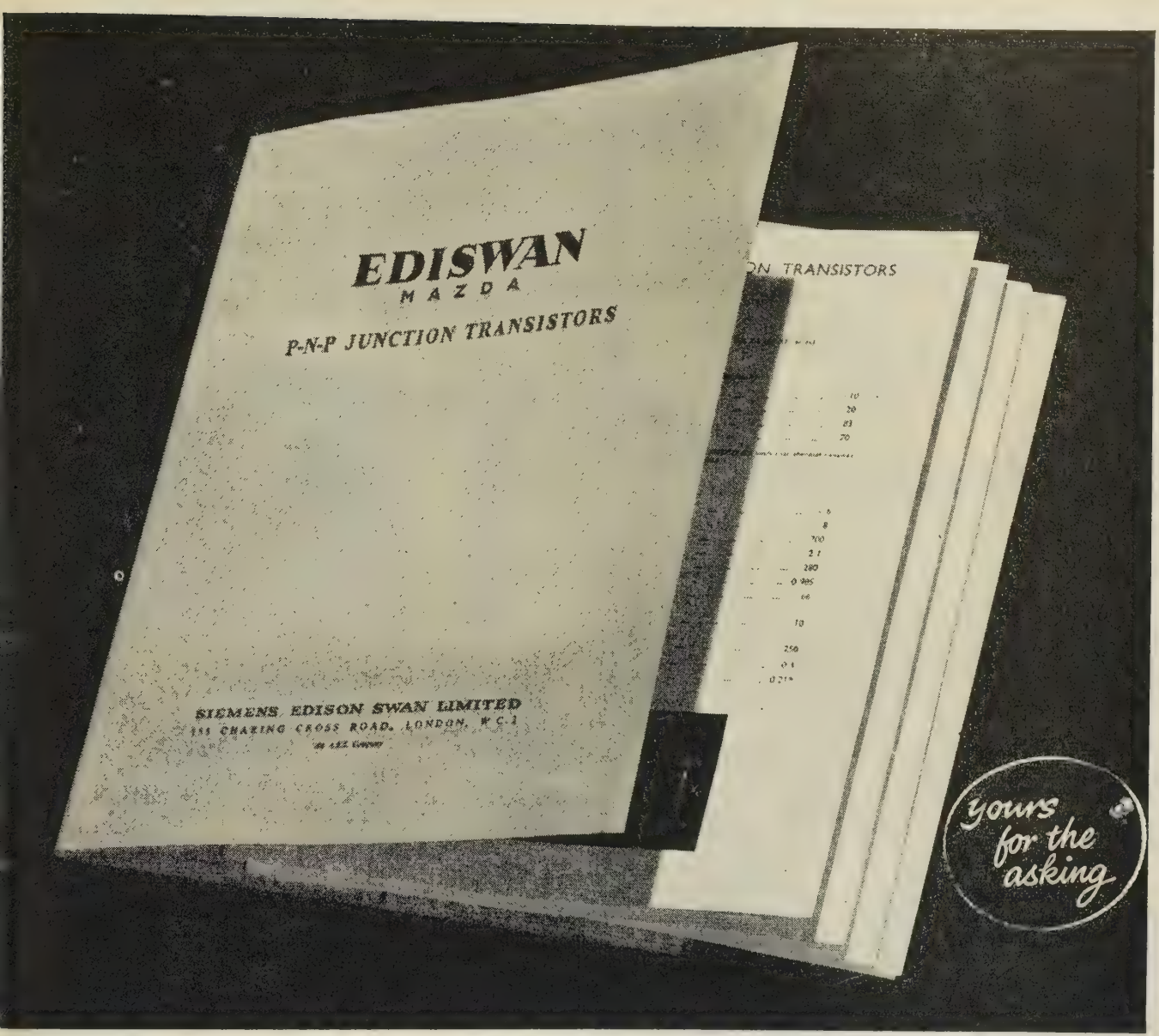
See us at the R.E.C.M.F. Exhibition

Stand 22
Grosvenor House
April 14th—17th

Mullard

'TICONAL' PERMANENT MAGNETS
MAGNADUR CERAMIC MAGNETS
FERROXCUBE MAGNETIC CORES

MULLARD LIMITED, COMPONENT DIVISION, MULLARD HOUSE, TORRINGTON PLACE, W.C.1
MC260



EDISWAN transistors

MAZDA

If you are manufacturing or designing electronic equipment you will find this folio of data sheets helpful as a source of reference. It gives you comprehensive information and characteristic curves covering the whole range of EDISWAN Mazda transistors.

Simply ask for the P-N-P Transistor Folio on your business notepaper.

at highly competitive prices

SIEMENS EDISON SWAN LIMITED

155 CHARING CROSS ROAD, LONDON, W.C.2. AND BRANCHES.
TELEPHONE: GERRARD 8660. AN A.E.I. COMPANY.
TELEGRAMS: SIESWAN WESTCENT, LONDON.

CRC 25/4

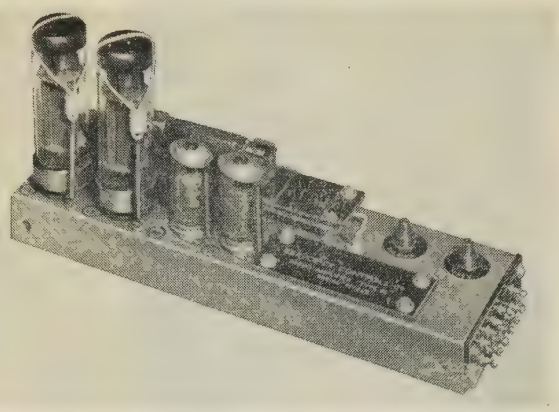
It's a SERVOMEX SA.61 Servo-amplifier, but look at the price . . .

£17

IDEAL FOR EXPERIMENTAL WORK & TEACHING

This new servo-amplifier is used as a standard sub-unit in our automatic control systems and is now available separately from stock at the very low price of £17.0.0. Suitable for most split-field d.c. servo-motors the amplifier has a gain of 2,000 mA per volt and uses components approved for operation up to 70 °C.

Please write for full specification.

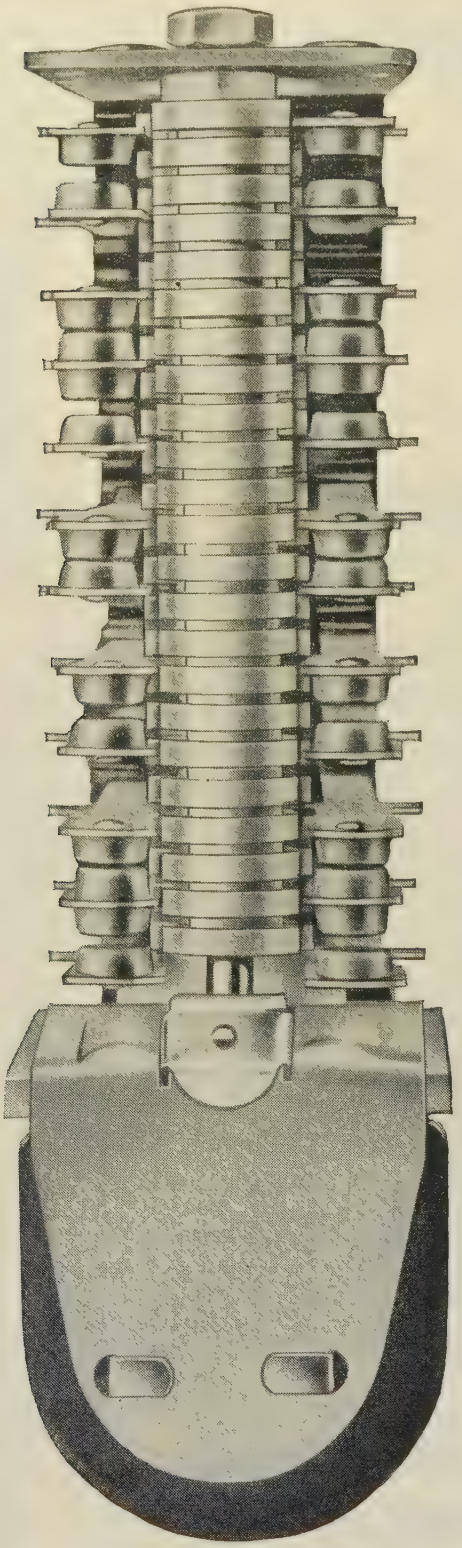


also look at the New Servomex Price List— price reductions beat price increases, 13-0

Our new price list effective January 1st shows no price increases. On the other hand it includes 13 price reductions! This in spite of increasing costs of raw materials.
Increased demand and still further improved production methods once again benefit all our customers.

Send for your copy

Servomex Controls Limited · Crowborough Hill · Jarvis Brook · Sussex · Crowborough 1247



You don't often see a relay like this!

Series 285

- **Fast operating**
- **Negligible bounce**
- **Exceptional life**
- **Exceptional insulation**

MAGNETIC DEVICES LIMITED

Magnetic Devices LTD.

A.I.D. AND A.R.B. APPROVED

MAGNETIC DEVICES LTD., EXNING ROAD, NEWMARKET, SUFFOLK

Phone: Newmarket 3181/2/3

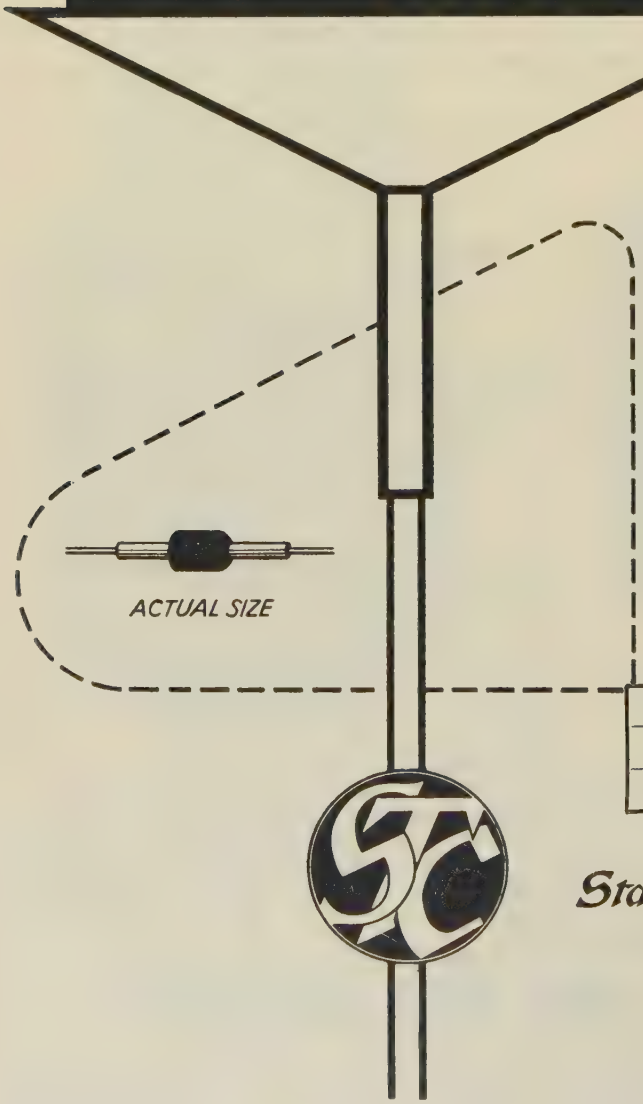
Grams: Magnetic Newmarket

MD 21

INDEX OF ADVERTISERS

Firm	page
Adcola Products Ltd.	xxxiv
Aero Research Ltd.	xxviii
Airmec Ltd.	xxxvi
Automatic Telephone & Electric Co. Ltd.	xxi
British Thomson-Houston Co. Ltd.	xxix
Cable Makers Association	xiv
Cathodeon Crystals Ltd.	
Cinema Television Ltd.	iv
Cossor Instruments Ltd.	xxxiii
Dewhurst & Partner Ltd.	xxvi
Dubilier Condenser Ltd.	xix
English Electric Valve Co. Ltd.	ii & iii
Erie Resistor Ltd.	xx
Ferranti Ltd.	xvii
P. X. Fox Ltd.	
General Electric Company Ltd. (Telecommunications)	vi & vii
R. Johnson & Nephew Ltd.	xl
Lodge Plugs Ltd.	xxxii
London Electric Wire Co. & Smiths Ltd.	xxxviii
Claude Lyons Ltd.	xliii
Magnetic Devices Ltd.	xxx & xlii
Marconi Instruments Ltd.	xvi
Marconi's Wireless Telegraph Ltd.	viii–xi
Metropolitan Plastics Ltd.	xxxiv
Metropolitan-Vickers Electrical Co. Ltd.	
Mullard Ltd. (Magnetic Materials)	xxxvii
Mullard Ltd. (Valves)	xv
Multicore Solders Ltd.	
Newmarket Transistor Co. Ltd.	xiii
Newton Bros. (Derby) Ltd.	xxxiv
Philips Electrical Ltd.	xxxviii
Plessey Co. Ltd. (A and A Group)	xii
Plessey Co. Ltd., Swindon	i
Salford Electrical Instruments Ltd.	xxxi
Savage Transformers Ltd.	xxvii
Servomex Controls Ltd.	xli
Siemens Edison Swan Ltd. (Exchange Equipment)	xxiv & xxv
Siemens Edison Swan Ltd. (Telecommunications)	xviii
Siemens Edison Swan Ltd. (Transistors)	xxxix
Siemens Edison Swan Ltd. (Telephone Apparatus)	xxii & xxiii
Solartron Electronic Group Ltd.	v
Standard Telephones & Cables Ltd.	xliv
Telephone Manufacturing Co. Ltd.	xxxvi
Texas Instruments Ltd.	xxxv
Westinghouse Brake & Signal Co. Ltd.	
Zenith Electric Ltd.	xl

SenTerCel
SILICON ZENER DIODES
Z 2 SERIES



are available from production

HAVE A LARGE DISSIPATION FOR THEIR SIZE

are available from production

ARE SUITABLE FOR HIGH TEMPERATURE OPERATION

are available from production

HAVE A LOW TEMPERATURE CO-EFFICIENT OF VOLTAGE

are available from production

ARE SUITABLE FOR USE AS REGULATORS, LIMITERS, SURGE SUPPRESSORS, AND REFERENCE VOLTAGE APPLICATIONS

are available from production

± 20% VOLTAGE RANGE					
ZENER DIODE TYPE	Z2A33	Z2A47	Z2A68	Z2A100	Z2A150
NOMINAL VOLTAGE	3·3	4·7	6·8	10·0	15·0

Standard Telephones and Cables Limited

Registered Office : Connaught House, Aldwych, London, W.C.2

RECTIFIER DIVISION

EDINBURGH WAY · HARLOW · ESSEX

THE PROCEEDINGS OF
HE INSTITUTION OF ELECTRICAL ENGINEERS

EDITED UNDER THE SUPERINTENDENCE OF W. K. BRASHER, C.B.E., M.A., M.I.E.E., SECRETARY

OL. 105. PART B. NO. 20. MARCH 1958

.396.946:621.396.969.35

The Institution of Electrical Engineers
Report No. 2552
Mar. 1958

RADIO OBSERVATIONS ON THE RUSSIAN SATELLITES

A Discussion Meeting

a meeting of the RADIO AND TELECOMMUNICATION SECTION, *on the 22nd November, 1957, a number of short contributions were presented, describing the part played by radio in observing the behaviour of the Russian earth satellites. There were two sessions, at each of which the contributions were followed by a discussion. At the end of the second session there was a general summing-up by* Dr. R. L. SMITH-ROSE.

ession I

INTRODUCTION

R. L. F. BOYD, Ph.D., B.Sc.(Eng.), Associate Member.

We at University College are, of course, very interested in the bject of artificial satellites because, with others, we are con-rned in the British research at high altitudes, using the excellent le Skylark rocket so successfully built for us by the R.A.E. is vehicle, it is hoped, will ultimately enable us to attain itudes of something like 300 km, and it is beyond this altitude at it really becomes worth considering the use of satellites for ploration.

First let us consider the fact of these launchings. I think that e time of the launchings came as no great surprise, but most of were surprised to learn the weight of material that the Russians d fired into the upper atmosphere. At the beginning of this eeting we ought to say that we regard this as a very con-lerable scientific achievement, and we compliment those sponsible on the result that they have achieved. There is le doubt that the degree of precision in the launchings is cellent, and quite above what might have been expected.

I would, however, go on to say, with due respect to the ussians, that this may not be the best way of doing research these altitudes. I believe that the American approach, with a nimal vehicle, the Vanguard, is in the long run—provided it es what it is intended to do—a better method of doing upper nospheric research, in that it is a great deal cheaper. On the ner hand, if there are, linked to a military programme, vehicles ich are capable of taking 1 ton to the upper atmosphere, we scientists are pleased to make use of them. There is little bt that there has been a considerable achievement in rocketry , and it is likely that the rocket motors are an advance on kind of thing intended for the Vanguard vehicle.

et us now consider the scientific purpose behind these chings. It is possible to impress the world at large tre-dously by sending very weighty vehicles to circulate round world, but, of course, the primary purpose in an experiment his kind is to obtain data on conditions in the upper atmo-re and in outer space. It is clear that the intention of the Russians in the launching of Satellite I was first to show that it is possible to launch a satellite of this kind; secondly, to study the ionosphere by means of the radio link between the satellite and the ground; thirdly, by observations on the orbit and on the perturbations of the orbit, to obtain data on the figure of the earth; and finally, since a spherical satellite was chosen, to obtain data of the drag of the atmosphere on the satellite itself.

On the other hand, Satellite II is similar to Satellite I in almost every respect, except that the sphere containing the transmitter has not been ejected from the rocket, so that there is a large and massive satellite circulating.

The problem of obtaining data on the upper atmospheric density by means of drag is complicated by the irregular shape of the second satellite, and it is also clear that one of the major intentions here is not a scientific project at all; it is a project in space exploration. The purpose is to discover something about the hazards of space travel. Probably the greatest hazard, or the greatest uncertainty so far as hazards are concerned, is that of cosmic rays. Since it is clear that the observations on the dog fired in Satellite II have not lasted long enough for an adequate study of their effect to be made, I suppose we can assume that the Russians have intended to discover first the effect on a mammal of the acceleration during launching, and then the effect of a prolonged sense of weightlessness. It is reported that the second satellite carries devices for measuring the incoming flux of cosmic rays and devices for measuring the flux of solar ultra-violet light.

The satellite is a particularly useful means of observing solar radiation, because of the importance of the sudden variations in the Lyman-α and X-radiation. It is difficult by means of rockets to observe these fluctuations satisfactorily, as one never knows when an outburst on the sun is going to take place. Efforts have been made to have rockets ready to be launched, either hanging from balloons at a great altitude or standing by on zero-length launchers to be fired at the first signal from an

observatory that a flare has occurred; but the satellite is there all the time, and that offers a great advantage for this type of study.

I wish to turn your attention to cosmic rays for a moment. In the second column of Table 1 are given, so far as they are

Table 1

COMPOSITION OF THE PRIMARY COSMIC RAYS AT GEOMAGNETIC LATITUDE 41·5° N

(from Webber)

Component	Intensity	Astronomical abundance
	%	%
H	<86	88
He	12	12
Li, Be, Bo	0·5	10^{-6}
Atomic No. ⩾6	1·8	0·12

known at present, the intensities of the cosmic rays having the composition given in the first column. There are two things here which are important. From the point of view of the space hazard—which I do not feel to be a tremendously important scientific investigation, but one which interests those concerned in space travel—the particles of large mass (in the last line of the Table) are those which are most damaging when they pass through the organs of the body because they are capable of producing intense ionization along their track. The question of importance is whether the rate of cause of damage to the organs by these intensely ionizing radiations is greater than the rate of repair that the body can carry out. Upon the answer to this question depends ultimately whether or not space travel for human beings is a possibility.

From the point of view, however, of our understanding of the universe in which we live the third line of the Table is of much greater significance. It will be noticed that the light particles—Li, Be and Bo—have fluxes in the atmosphere which are greatly in excess of the ordinary astronomical abundance of these substances, given in the last column. The significance of this is that these particles arise from the fragmentation of heavier particles, and from a measure of their relative abundance we can hope to obtain information concerning the path which the cosmic rays have travelled from their point of origin in the universe.

Table 2 is a list of some of the studies for which the satellite

Table 2

SATELLITE STUDIES

Geophysical	Celestial
Density, temperature and composition above 200 km	Solar (and stellar) spectroscopy
Ionization beyond maximum of F-layer; radio propagation	Cosmic rays—Heavy component. Flux of Li, Be, Bo
Geodesy	Solar storm phenomena
Geomagnetism	Meteors and interplanetary dust
Meteorology and albedo of the earth	Interplanetary gas

is particularly suitable. In the first column are the geophysical studies. Of primary importance are the data which can be obtained by means of the satellite on the density, temperature and composition of the atmosphere at heights greater tha 200 km. It becomes increasingly difficult at these heights t make measurements by means of rockets. The gas carried u by the rocket and by the instruments by which such measuremen are made is a serious difficulty. The pressure in any case is s low that it is much better than the best vacuum normally use in the laboratory.

Probably of comparable importance is the study of th ionization beyond the peak of the F-layer. The ordinary metho of radio sounding depend on reflection, and beyond the peak c the F-layer appreciable reflection does not take place, so tha it is very difficult to study the ionization density from the groun It is not true to say that we can get no data. We can refle signals from the moon, and we can make use of radio-astronom but the ordinary method is not suitable here.

The next study concerns finding the shape of the earth and th distribution of land masses, and so on. We can discover suc things as the size of the equatorial bulge, and eventually mak precise observation of the distances between the continent The problem of geomagnetism is also one on which the satellit can help, partly by giving an accurate overall survey of cosmic ray intensities. Cosmic rays are subjected to a magnetic fiel differing from that to be inferred by observations at the surfac of the earth, because of the currents flowing in the higher level of the upper atmosphere.

Then we have the study of meteorology, the exploration b photographic or phototelemetric methods of the distribution c clouds, the gathering of hurricanes and phenomena of this typ and also of the amount of energy reflected from the surface of th earth.

In the celestial column there is the whole subject of sola and stellar spectroscopy, in regions of the spectrum where ligh is absorbed by the atmosphere itself. I have already mentione the cosmic-ray flux and the solar-storm phenomena. There i the possibility of obtaining information on the density of th micro-meteoric dust, regarding which there is at present a dis crepancy between the various methods of observation amountin to a factor of over 1 000. Depending on the character of th dust and upon the quantity coming in, this may possibly hav an effect on the weather. There seem to be some data—thoug this is disputed—to show that there is a displaced correlatio between the meteoric showers and the days of wettest weather.

There is the further possibility of determining the concentratio of interplanetary matter by means of its fluorescence in the ray of the sun.

Table 3 shows the forces acting on the satellite. There are th gravitational forces, the geocentric force, the forces due to th non-uniform distribution of matter, and the forces due to th

Table 3

FORCES ACTING ON ARTIFICIAL SATELLITES

Gravitational	Electromagnetic	Aerodynamic
Geocentric force	Eddy current damping of spin	Airdrag
Non-centric geophysical forces (equatorial bulge, mountain masses, non-uniform density distribution)		Westerly wind of rotation Bernoulli forces Air damping of spin
Celestial gravitation (sun, moon)	Electrostatic drag	

n and moon and other celestial bodies. The effect of the sun d moon is very small, and the perturbations become appre- able only over hundreds of years.

There is electromagnetic damping of the rotation of the satellite ue to eddy currents induced in the earth's field, and also an ectrostatic drag component due to the fact that the satellite comes charged to a potential of some ten volts or more by st electrons and drags after it a space-charge of ions of the pposite sign which augment the drag of the upper atmosphere the vehicle. This drag is comparable, at 300 miles, with that erted by the atmospheric particles themselves.

The fact that the earth is rotating beneath the orbit of the tellite, and that there is therefore continuously a westerly nd acting on that orbit, would eventually cause detectable rturbation. There are also the Bernoulli forces on the satellite ue to the fact that it is rotating and that the pressures on the ont and the back of the satellite are not the same.

In conclusion, I will point out the principal elements of the orbit that it is important to determine. The observations of the earth's satellite are aimed at determining six quantities. If we can determine these and their variation with time we can explore with an otherwise inert satellite the distribution of mass in the earth and the effect of air drag and so on. These quantities are:

(i) The height of the perigee.
(ii) The height of the apogee.
(iii) The position in the celestial sphere of the ascending node, i.e. the intersection of the orbit plane with the equatorial plane on the side of south-to-north passage.
(iv) The angle which the line joining the apogee to the perigee makes with the ascending node.
(v) The inclination of the orbit, i.e. the angle between the orbital plane and the equatorial plane.
(vi) The time of passing the ascending node.

It is part of the business of this meeting to discuss how radio measurements assist in the determination of these quantities.

OBSERVATIONS AT CAMBRIDGE

J. R. SHAKESHAFT, M.A., Ph.D.

The news of the first satellite reached the Cavendish Laboratory rly on Saturday, 5th October, the satellite having been launched e previous evening. It was decided to adapt existing equipment order to make observations of its track. We hoped that such bservations might be of use in predicting positions for optical d radar observers, because, of course, the inherent accuracy optical observations is very much greater than can be achieved radio measurements. In fact, however, the climate of this and being what it is, few photographs have been taken, and I lieve that I am correct in saying that much of the work done the Nautical Almanac Office has been based on radio obser- tions.

On the night of 5th–6th October it was found that a strong gnal was detectable from the satellite, and on the 6th an inter- rometric aerial was set up at the Mullard Radio Astronomy bservatory. This consisted simply of two horizontal full-wave poles at 40 Mc/s, separated along an east-west line by a spacing four wavelengths, which at 40 Mc/s is just under 100 ft. The equencies of transmission were 40·002 and 20·005 Mc/s. At e start, when the first satellite went up, the signals were inter- ittent, with a repetition frequency of about 3 per sec. We are ld that this pulse frequency was a measure of the temperature side the satellite.

The higher frequency was chosen for the interferometer in der to reduce ionospheric refraction effects, though observa- ons were made later at 20 Mc/s and also at the second harmonic the 40 Mc/s transmission to learn about these effects.

The two 40 Mc/s aerials making the interferometer were con- ected to a 'phase-switching' receiver, the output of which ables the phase difference between the signals arriving at the o aerials to be determined at short intervals. The reception ttern of the interferometer is a series of fringes which intersect y horizontal plane above the aerials in a set of hyperbolae ich are loci of constant phase difference, their shape depending the height of the plane. To a first approximation the satellite thin, say, 500 km of the observing station may be considered moving in such a plane.

Fig. 1 gives some idea of the orbit of the satellite in relation this country. The inclination of the orbit which Dr. Boyd entioned, the angle which the plane makes with the equatorial ane, is nearly 65°, which means that the maximum latitude

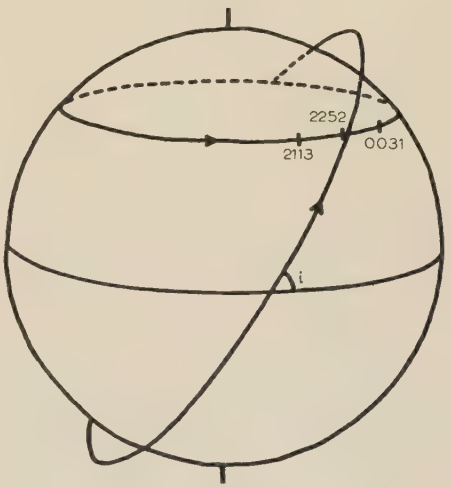


Fig. 1.—The orbit of Satellite I in relation to the latitude of Cambridge.

The times indicate the approximate position of Cambridge during the successive transits.

reached by the satellite is also 65°. The small circle on which Cambridge lies, at a latitude of 52°12′, is marked. We can consider the orbital plane as fixed, with the earth rotating so that Cambridge moves eastwards along the small circle beneath the orbit. The position of Cambridge at successive times of transit is marked. It will be seen that, since the satellite can be detected at 40 Mc/s within a range of about 2 000 km, 7 transits could be observed out of the total of 15 in a day. These 7 in fact occurred for the first satellite during the night. This was fortunate, in spite of the discomfort of night work, because it meant that the severe daytime interference which we have been experiencing and which upset the recording of the second satellite was avoided.

A rough value of the inclination was found from the rate of transit of the satellite across the interference fringes when the satellite was passing far to the north of Cambridge. The satellite was then moving parallel to the line of the aerials and transversely

across the fringes. The velocity used was initially calculated from the measured period by assuming the satellite to have a circular orbit. From the inclination, the angle at which the track crossed the meridian at Cambridge was found and it was then possible, as shown in Fig. 2, to use the observed times of zero output to

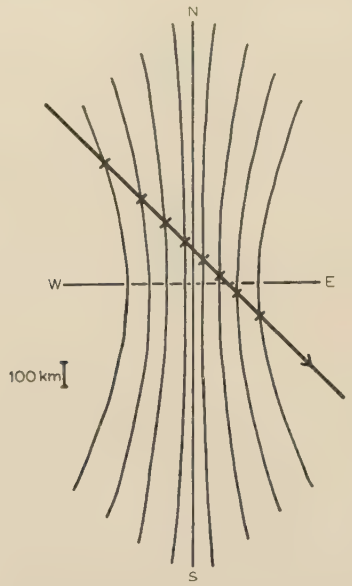


Fig. 2.—Grid lines for a height of 530 km showing the derivation of the track relative to Cambridge.

determine a unique track and height. The times were converted into distances by means of an assumed velocity and then the track was adjusted, maintaining the correct orientation, until these positions coincided with the hyperbolae. The time of crossing the meridian could then be read off, and also the location of the track with respect to the aerials.

For the more distant tracks, account had to be taken of the curvature of the orbit, and when an elliptical orbit had been derived, the variation of velocity had to be considered.

Measurements of the apparent changes of frequency due to the Doppler effect were started on the 6th October, but I will not dwell on the details as other speakers will be describing them.

Table 4

Inclination	64°40′ ± 10′
Period between successive northward crossings of lat. 52°12′	5 750·0 solar sec ± 0·3 sec
Rate of decrease of period	2·2 ± 0·1 sec/day
Height at latitude 52°12′ on northward passage	223 ± 10 km
Height at latitude on 52°12′ southward passage	470 ± 15 km
Semi-major axis	6 937 km
Eccentricity	0·053 ± 0·001
Latitude of perigee	36° N. ± 3°
Longitude of ascending node	308°09′ ± 15′
Retrograde motion of ascending node	3°10′ ± 6′
Meridian passage at Cambridge on 15.10.57	21 h 8 m 41 s U.T.
Height above mean radius of earth at perigee	197 ± 10 km
Height above mean radius of earth at apogee	934 ± 10 km

Our Doppler measurements were combined with those from other stations, and also with the interferometer results, to derive an approximate orbit for the 15th October (Table 4).

The decrease in period of about 2 sec per day corresponds to a change in the semi-major axis of 3·5 km per day. It can be shown that, for an elliptical orbit such as this, all but a few percent of the change occurs on the height at apogee, the greatest height, so that the height at perigee hardly alters.

Because of the importance of the rate of change of period in connection with the determination of the air density, the variation of the period is shown plotted in Fig. 3. A curvature is detectable

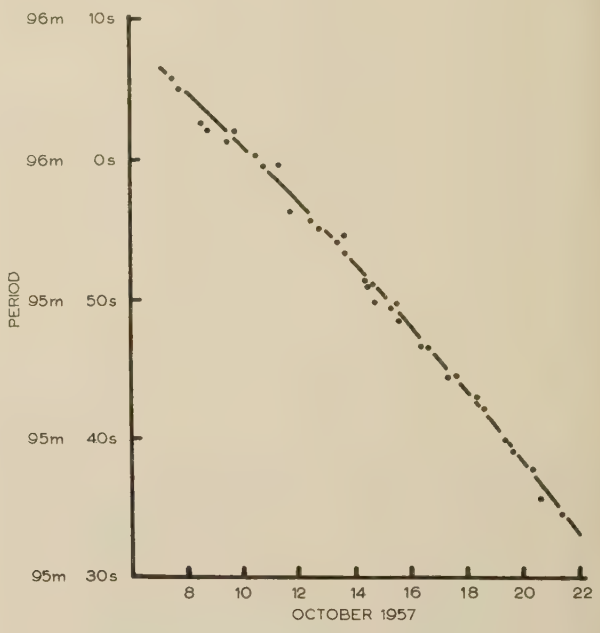


Fig. 3.—Variation of the true period.

showing that the satellite is getting down to regions of greater density. By assuming a mass of 84 kg and a diameter of 58 cm for the satellite, as stated by the Russians, and an atmospheric temperature of 1 000° K, an approximate value for the air density at 200 km is obtained of 4×10^{-13} g/cm^3. This calculation ignores any electrical drag effects. A more accurate determination will no doubt become available in due course.

At the risk of trespassing on other speakers' preserves, I must mention the fascinating periodic fluctuations of intensity which have been observed at all three frequencies. We believe these to be caused by a combination of the rotation of the satellite itself at about 7 r.p.m. and the Faraday effect, i.e. the rotation of the plane of polarization of the transmitted wave in the ionosphere due to the earth's magnetic field. A detailed examination of these effects, and also of the Doppler records, is in progress at the Cavendish Laboratory to investigate the ionosphere, but no results are yet available.

The transmitter in the first satellite lasted three weeks, but that in the second only five days. The orbit of the second satellite is closest to Cambridge during dawn and daylight hours, and daytime interference spoilt many observations. The orbit, however, is very similar to that of the first satellite, except that it has a larger major axis and is more elliptical, the lowest height being about the same as before but the highest about 1 600 km instead of 900 km. The period is 103 min.

I should like to end by thanking all those workers in other laboratories who have helped us so much during the last seven weeks.

APPARATUS USED AT THE ROYAL AIRCRAFT ESTABLISHMENT

A. N. BERESFORD, B.Sc.

I am going to describe the apparatus used by the Radio Department at the Royal Aircraft Establishment, with which we tracked the two satellites. I shall start with the characteristics of a germanium junction diode (the T19G), which, overall, is 5 mm long and has a diameter of 3 mm. From the d.c. characteristic (Fig. 4) it will be seen that 0·8 volt drives about 100 mA through the diode and that the slope resistance is considerably less than the d.c. resistance.

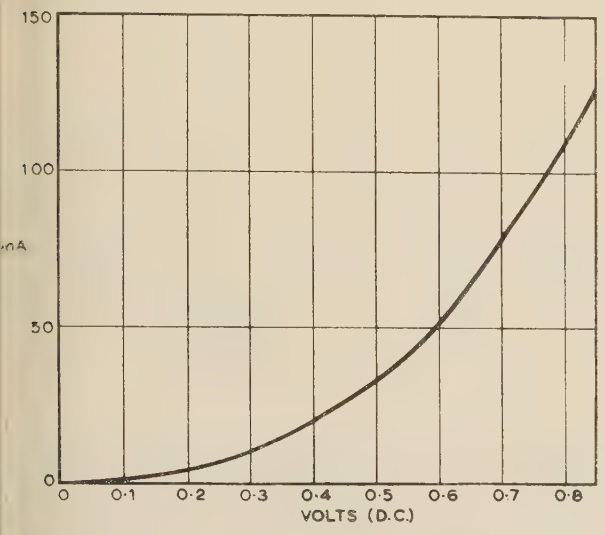


Fig. 4.—D.C. characteristic of germanium junction diode.

We have made r.f. measurements on this diode with the apparatus shown in Fig. 5(*a*). The technique, which will be familiar to most of you, is particularly well suited to this type of measurement: varying direct currents can be sent through the diode from the input end of the slotted line, chokes and condensers at this end having no effect on the measurement. With l.f. modulation on the signal generator, and an a.f. amplifier following the detector, measurements can be made with an input of about 0·1 volt (r.m.s.).

The graph [Fig. 5(*b*)] shows the r.f. resistance at 200 Mc/s that is obtained by these means for various direct currents. Above about 50 mA, the r.f. slope resistance varies little and is about 3 ohms. On this graph are also shown, as circles, the slope resistance for the same crystal obtained by drawing tangents to the d.c. characteristic: some correlation is evident. Reversing the battery in Fig. 5(*a*), the cut-off impedance is measured and is found to be purely reactive, corresponding to about 0·6 pF. This value is very nearly constant with changes in direct voltage from zero to −12 volts.

This diode, then, is ideally suited to function as a small-signal switch, being equivalent to 3 ohms when cut on and 0·6 pF when cut off. This is a switching ratio at 40 Mc/s of more than 1 000 : 1. Before describing how this switching facility is used in the interferometer, I should like you to consider Fig. 6, which

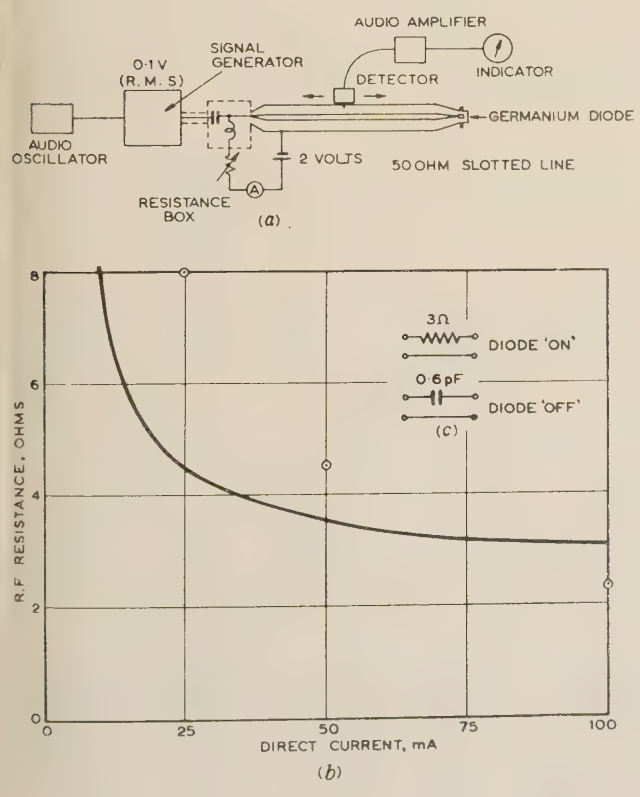


Fig. 5.—200 Mc/s measurements.

(*a*) Schematic.
(*b*) Relation between r.f. resistance and direct current.
○ D.C. slope resistance.
(*c*) Equivalent circuit.

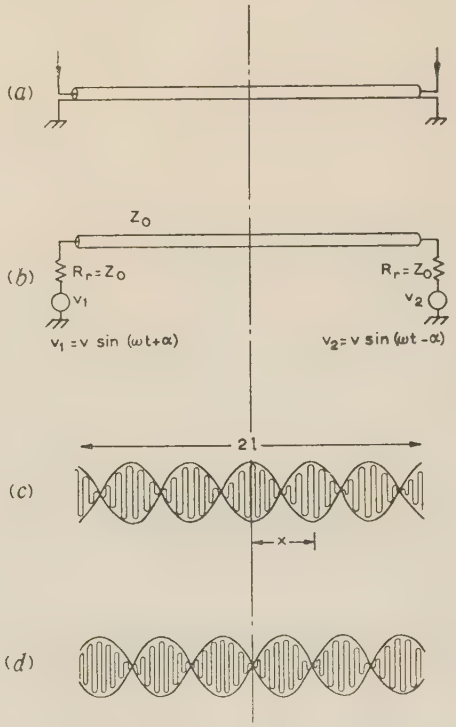


Fig. 6.—Two identical aerials connected via a coaxial cable.

(*a*) Interconnection of two aerials.
(*b*) Equivalent circuit.
(*c*) V.S.W.R. pattern for $\alpha = 0°$.
(*d*) V.S.W.R. pattern for $\alpha = \pm 90°$.

shows two identical aerials interconnected by a coaxial cable. Three simplifying assumptions will be made:

(i) That each aerial has a radiation resistance equal to the characteristic impedance of the cable.

(ii) That the signals picked up by the two aerials have the same amplitude, but not necessarily the same phase.
(iii) That the cable is lossless.

Under these conditions we have the equivalent circuit as shown at (*b*). An infinite v.s.w.r. will be established in the cable given by

$$V = v \sin\left(\omega t - \frac{2\pi l}{\lambda_c}\right) \cos\left(\alpha - \frac{2\pi x}{\lambda_c}\right)$$

where $2l$ is the length of cable, x is measured from the cable centre and λ_c is the wavelength in the cable. That is, an amplitude pattern is set up in the cable and the phase of this pattern is directly proportional to the phase difference appearing at the aerials.

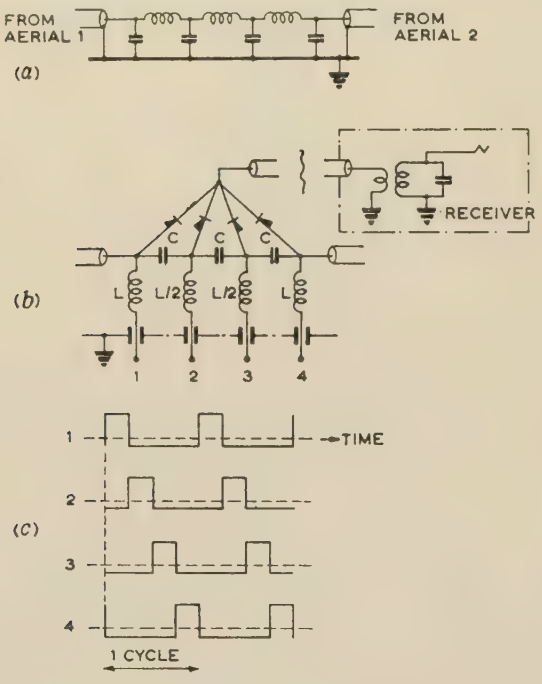


Fig. 7.—Radio-frequency crystal switch.
(*a*) Normal form of artificial line.
(*b*) Artificial line actually used.
(*c*) Switching waveforms: amplitude = 1 volt, peak to peak.

If we now break the cable at the centre and insert an artificial line having the same characteristic impedance as the cable, the pattern is not disturbed, but we can now feed the pattern into a receiver as amplitude modulation of the incoming carrier (Fig. 7). The form of artificial line actually used is shown at (*b*). The electrical length of the line is 135° with the germanium diodes at 45° intervals. This equivalent network, (*b*), is more convenient than (*a*) as it enables the diodes to be switched via the shunt inductances.

It would, of course, be perfectly possible to use a mechanical switch instead of the diodes, but the latter have the following advantages:

(i) The switching speed may be varied over wide limits—from, say, 5 c/s to 10 kc/s.
(ii) The change-over from one diode to the next may readily be made faster than 1 microsec.
(iii) The unit is extremely compact—at 40 Mc/s about twice the size of a matchbox.

Against these advantages must be set the complexity of the electronic circuits required to produce the four switching waveforms, (*c*). The requirements here are an extremely-low-impedance source—we are switching about 8 ohms—and exact similarity of the four waveforms except for the time displacement. By this means a modulated radio frequency of the form shown in Fig. 8(*a*) is fed into the receiver. Before proceeding further one or two points need clarification.

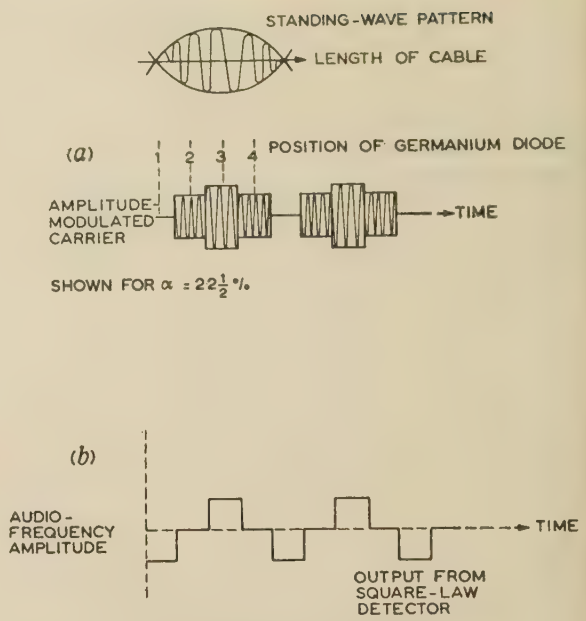


Fig. 8.—Modulation pattern.
(*a*) Radio-frequency envelope.
(*b*) Audio-frequency output.

For optimum power transfer the receiver should have an input impedance of half the cable impedance—25 ohms in a 50-ohm system. During the time that any one crystal is 'on', therefore, this impedance shunts the transmission line. Provided, however, that this impedance is the same for all four diodes (and this is

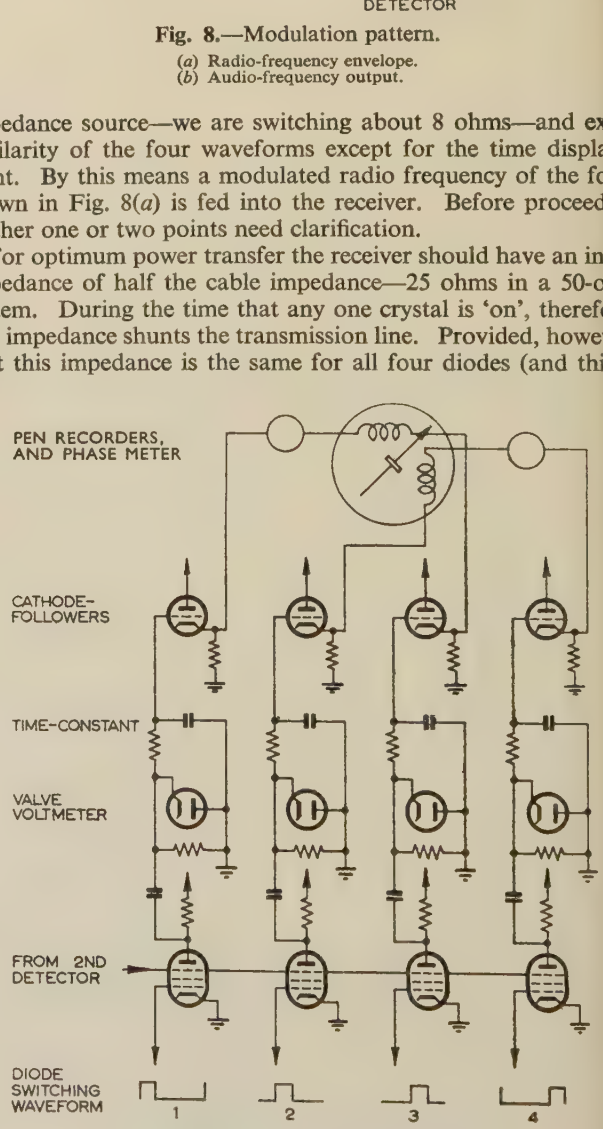


Fig. 9.—Phase-comparison circuit.

n series with the 3 ohms odd of the diode), it may readily be hown that no distortion of the standing-wave pattern results, ut it should be clear that the lower the diode slope resistance, he more easily will this condition be met. A further point vorth noting is that, from the junction of the four diodes nwards, the system is quite independent of radio-frequency hase—the phase-sensitive role of the interferometer is ccomplished.

The diode switch having converted the spare standing-wave attern into a signal modulated in time, the original equation ay be rewritten

$$V = v \sin\left(\omega t - \frac{2\pi l}{\lambda_c}\right) f(\cos \overline{\alpha - \omega_s t})$$

here ω_s is the angular switching frequency. This equation is imilar in form to that of a double-sideband suppressed-carrier ignal. To restore the original modulation, a square-law etector is therefore needed, giving a second detector output as n Fig. 8(*b*).

The phase of this audio output with respect to any of the riginal diode-switching waveforms is simply $2\alpha +$ (a constant), nd 2α, it will be remembered, is the r.f. phase difference between he signals in the two aerials. The simplest way of doing this phase comparison is to feed this audio output into two pairs of phase-sensitive amplifiers and rectifiers as in Fig. 9. The difference between output 1 and 3 is found to be simply $A \sin (2\alpha + 45°)$, and between 2 and 4, $A \sin (2\alpha - 45°)$. A continuous presentation of phase may then be obtained by applying these output direct currents to the coils of a phase meter. The angle indicated is then

$$\text{arc tan} \frac{A \sin (2\alpha + 45°)}{A \sin (2\alpha - 45°)} = 2\alpha + (\text{a constant})$$

In the interferometer these functions were also recorded independently on a two-pen recorder.

A word now about the aerials, which have been specially developed for this type of interferometer by W. T. Blackband. Each aerial consists of a 'bent' half-wave dipole (Fig. 10); this aerial has a horizontal polar diagram, as shown at (*b*), i.e. it is 'all-round looking'. The aerials were mounted $\lambda/4$ above ground as shown at (*c*), giving a degree of vertical directivity.

Features of this aerial are as follows:

(i) The radiation resistance, which may be varied by adjustment of the angle, is about 50 ohms for 90°.
(ii) The reactance can be made zero by alteration to the arm lengths and is approximately zero at $\lambda/4$.

These two adjustments are, to a first-order, independent. It will be remembered that correct matching to the interconnecting feeder is an essential part of the system.

For the first satellite, four aerials were assembled in two pairs

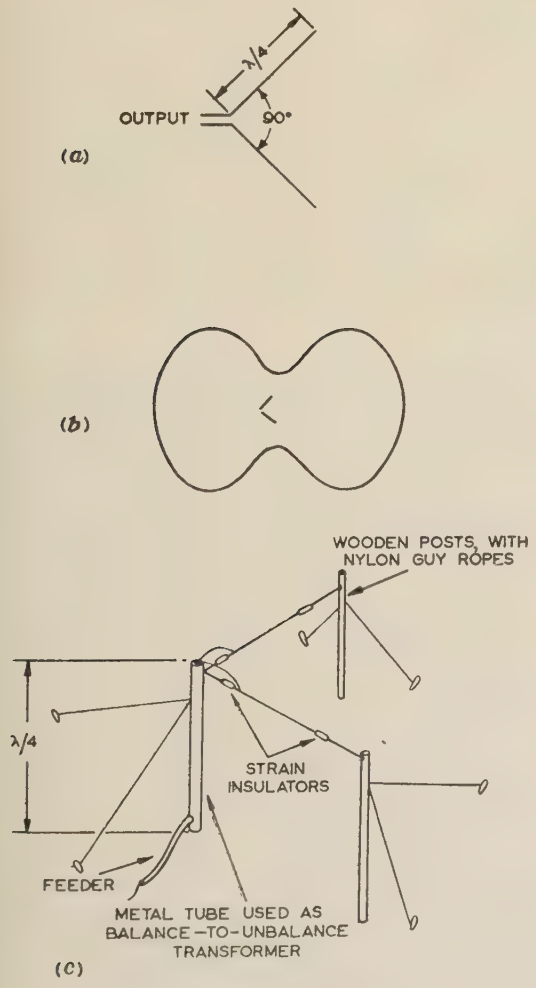


Fig. 10.—Aerial system.
(*a*) Folded half-wave dipole.
(*b*) Horizontal polar diagram.
(*c*) Layout.

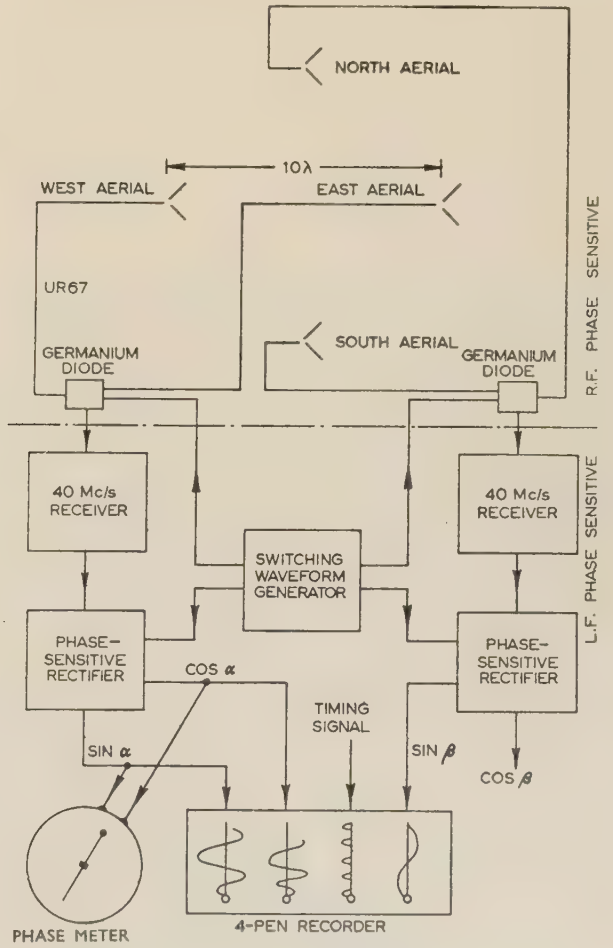


Fig. 11.—Aerial system.

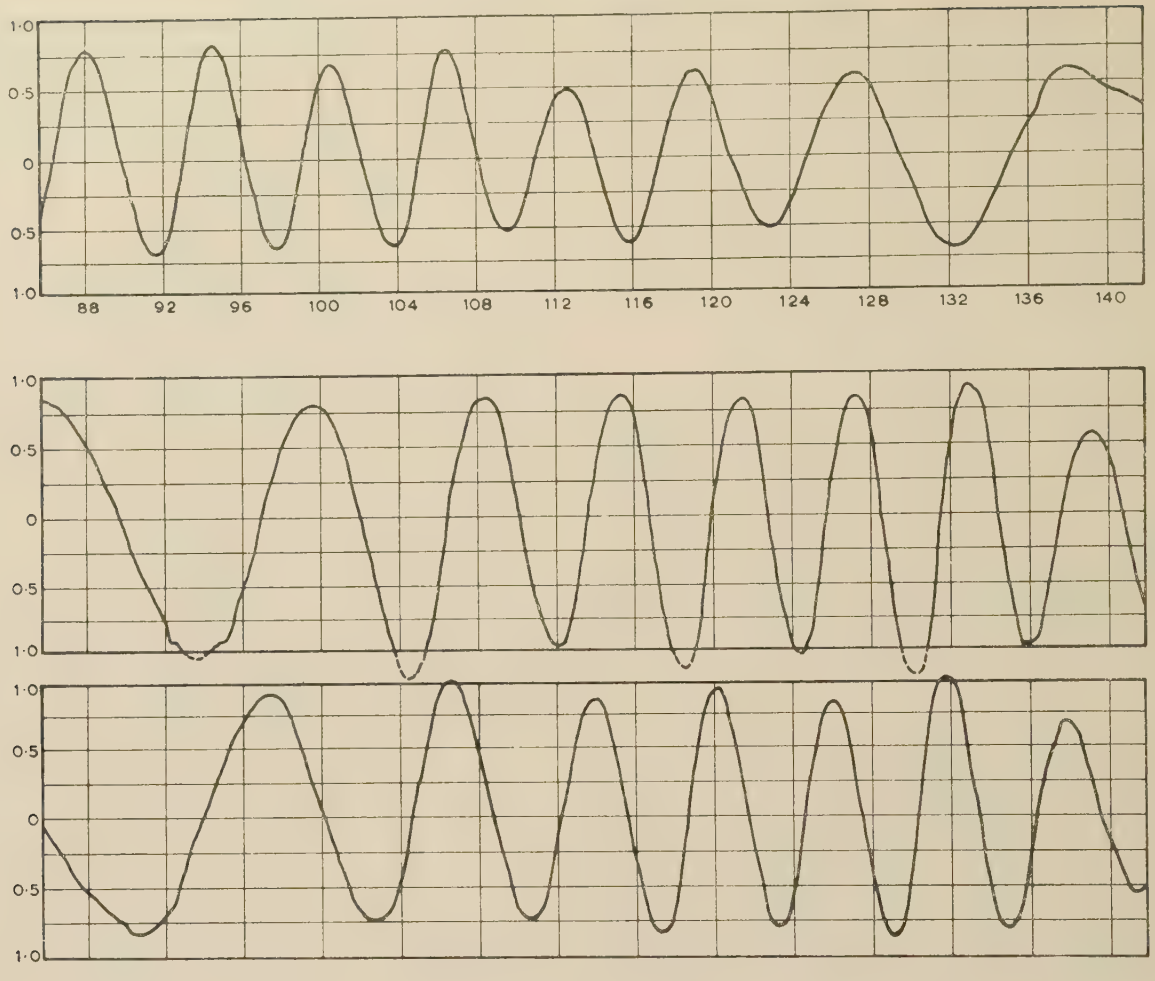


Fig. 12.—A typical recording.

placed accurately at right angles, and spaced at 4λ apart. For the second satellite this spacing was increased to 10λ. The complete system is shown schematically in Fig. 11 as it was assembled for tracking the satellites.

This scheme differs somewhat from the idealized one that I have been describing. Existing receivers with linear detectors had to be used, and only one four-pen recorder was immediately available—two pens for sine and cosine on one pair of aerials, the sine only on the other pair of aerials, and the fourth pen for the timing signal. Fig. 12 shows the type of recording obtained. The zeros are independent of the detector law, and the error introduced off zero can be calibrated out. The accuracy of the phase measurement is estimated at $\pm 3°$ of phase for the 50% error, corresponding to ± 3 min of arc of elevation on the equatorial plane. The interval between zeros is considerably more accurate than this, as it depends only on equipment stability during a transit.

In all, over 50 transits of the two satellites were recorded; the decoding of these records is the subject of a separate lecture.

In conclusion I should like to express my thanks to Dr. McPetrie and to Mr. W. A. Johnson, under whose guidance this development has been carried out. The design, construction and testing of the apparatus is largely the work of Mr. D. F. Fance, assisted by Mr. H. D. Mayhew and more recently by Mr. C. R. N. Barnard, all of whom are members of the Radio Department, R.A.E.

SOME DIRECTION-FINDING OBSERVATIONS ON THE 20 Mc/s SIGNAL

A. KITCHEN, B.Sc., Associate Member, E. R. BILLAM, M.Sc., W. R. R. JOY, B.Sc.(Eng.), Graduate, R. F. CLEAVER, B.Sc., Associate Member, D. L. COOPER-JONES, B.Sc.(Eng.), Associate Member, and J. M. BEUKERS, B.Sc.(Eng.), Graduate.

The provision of transmissions on a frequency of about 20 Mc/s om the Russian artificial earth satellites presented a unique pportunity for an investigation of ionospheric wave-propagation henomena, using direction-finding techniques. The equipment ed was of the type described by Earp and Godfrey[1] in 1947. principle, its operation is equivalent to the rotation of a mple vertical dipole round the circumference of a circle of out 7 wavelengths' diameter at 20 Mc/s. This rotation imposes phase modulation on the signal, the phase of which is used to rive the bearing information. One form of bearing display rovided a permanent record of signal bearing, signal strength, me reference and calibration data. Elevation could be derived rectly from the total phase excursion across the array.

The use of data of this type, in conjunction with simultaneous cordings of the Doppler frequency shift, is sufficient to establish tellite track information for local orbits. Plan range and slant nge at the instant of closest approach can be derived from the te of change of azimuth and of Doppler frequency, respectively, ing a value for satellite velocity obtained from the overall oppler shift on an overhead transit.

Some samples of the basic data obtained are shown in the companying illustrations. Fig. 13 refers to the evening over- ead transit on the 16th October. The bearing calibration is in eps of 10°, and the time markers are at 1 min intervals. It is a irly smooth transit, but some small irregularities will be noted. g. 14 shows another evening transit on the 17th October. It n be seen from the signal amplitude record that, during the cal part of the transit, the fading rate changed from about 13 6·5 per min. These fading rates can both be ascribed to inning of the satellite.[2] Fading at the lower rate is closely rrelated with some of the corresponding bearing scintillations.

We could obtain continuous data on the first satellite over nges up to several thousand miles, as shown in Fig. 15 for the me transit on the 17th October. The signal was heard for a uch longer time on the approach than on the recession. The rve at the top is the associated Doppler curve, which shows veral interesting variations. No systematic observations were ade on the early morning transits of the first satellite, from .W. to S.E., to see whether or not the asymmetry the reception times of the signal with reference the transit point was due to the normal ionization stribution in the F-region. Having regard to the w power of the transmitter, which we estimated be of the order of 1 watt, it is interesting to eculate whether the long-distance transmissions re received by one of the following mechanisms:

(*a*) Conventional multi-hop propagation between the earth's surface and the ionosphere (recalling that, at the relevant times, ionospheric absorption was negligible on the all-dark transmission path, and the earth reflections were from the sea).

(*b*) Propagation between the boundaries of the E- and F-regions, acting in the form of a leaky waveguide excited by the satellite transmitter itself.

Similar results were not obtained on the second tellite, the transmissions from which were very uch weaker. However, the local transits occurred edominantly in daylight over this country, when -region ionization densities were relatively high. hus, at times in the orbit when the satellite was ove the F-region, the 20-Mc/s transmission would efracted upwards into space. It was unlikely

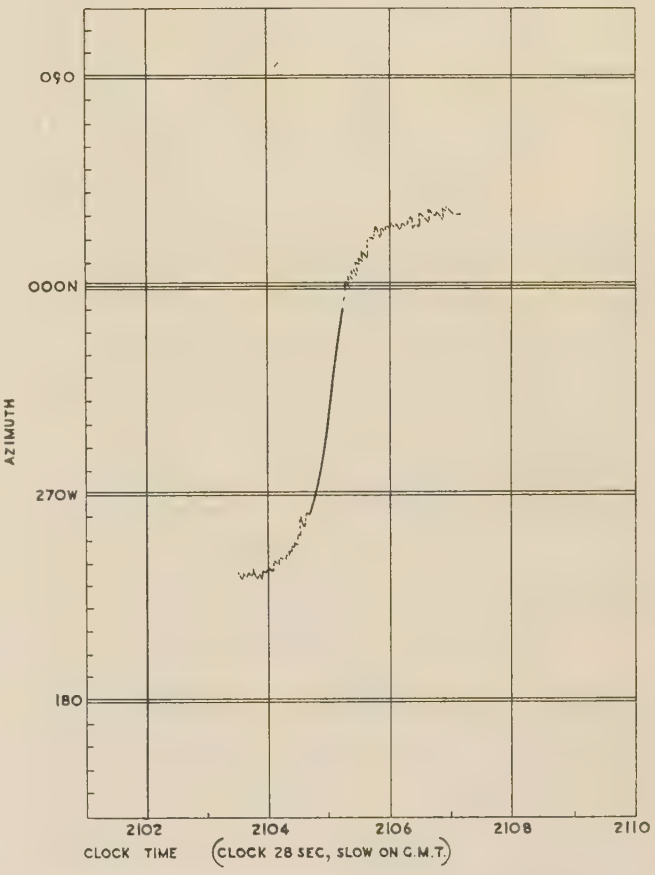


Fig. 13.—Bearing observations, evening overhead transit, 16th October, 1957.

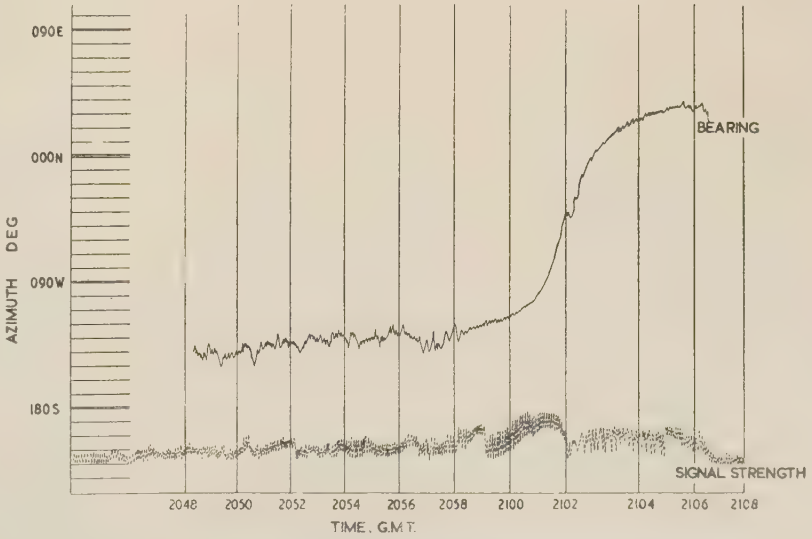


Fig. 14.—Signal strength and bearing observations, evening overhead transit, 17th October, 1957.

Table 5

SUMMARY OF LOCAL SATELLITE TRACK DATA. Velocity, 7·8 km/sec

Satellite	Date	Nominal time (G.M.T.)	Estimated track (true)	Height (nautical miles)	Time of crossing 51° 10.75′ N
I	15 Oct.	2100	042° ± 2°	143 ± 5	2109 h 43 ± 2 s
I	16 Oct.	2100	043.5° ± 1·5°	133 ± 3	2105 h 25·5 ± 2 s
I	17 Oct.	2100	044·5° ± 1·5°	Uncertain	2101 h 38·5 ± 2 s
II	6 Nov.	0640	045·5° ± 1·5°	152 ± 5	0637 h 36 ± 3 s

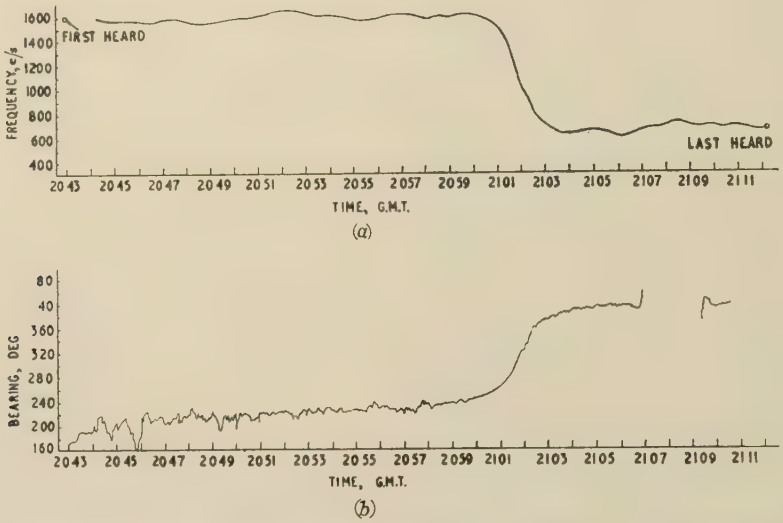


Fig. 15.—Doppler frequency and long-distance bearing observations, evening overhead transit, 17th October, 1957.

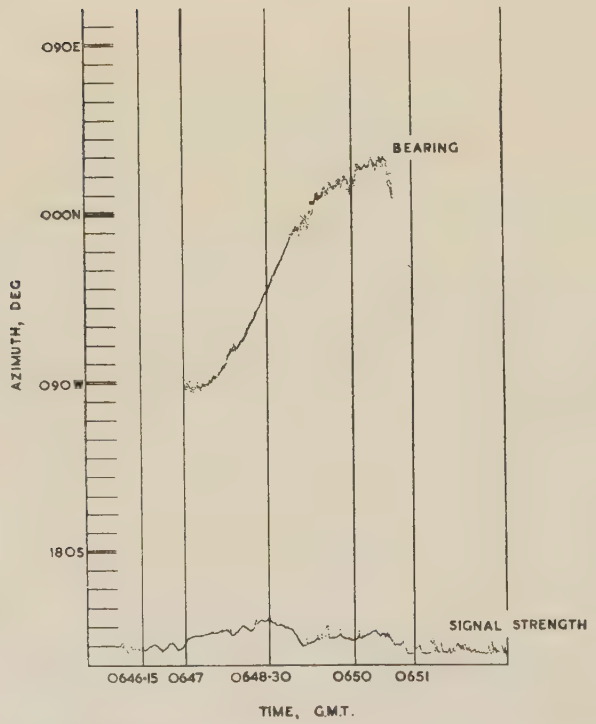


Fig. 16.—Bearing observations, morning overhead transit of second satellite, 7th November, 1957.

that a waveguide mode would be excited under these conditions. Further, at times when it was below the maximum in the F-region, the transmissions would be absorbed in the lower ionosphere owing to the seasonal rise in ionization density. The fact that the 40-Mc/s signals on a conventional receiver were stronger from the second satellite than from the first supports this.

An interesting feature of the simultaneous bearing and Doppler records is that the points of inflexion of the two, corresponding to the instant of closest approach, frequently differed by a few seconds. Initially, this was thought to be because the path of the orbit was not parallel to the surface of the earth. For some transits the apparent angle of climb of the first satellite as it went over us was of the order of 4°.

The bearing, signal strength, and Doppler records show many interesting variations about the mean values. Fig. 16 is a record of the bearing transit of the second satellite, on the 7th November. A large 'step' will be noted at one point, just before 0648 h. We have not yet had opportunity to study such effects in detail, but they are thought to be ionospheric in origin.

Provisional deductions from some of our measurements are given in Table 5. These seem to agree reasonably well with those published recently by the Cavendish Laboratory,[2] the R.A.E.[3] and others. However, we notice that on some of the more remote satellite transits—i.e. when the closest approach was a few hundred miles away—the effect of ionospheric refraction on the incident-wave paths appeared to reduce the slope of the bearing/time curve in the rapidly changing region, so that the range which we deduced from these data was somewhat in error.

The possibility arises that if we can have accurate information on the tracks from some other source to compare with our observations we may be able to deduce some facts about the structure of the ionosphere.

The work described was carried out as a joint programme by the staffs of the Admiralty Signal and Radar Establishment and Standard Telephones and Cables Ltd.

This note is published by permission of the Admiralty.

References

(1) Earp, C. W., and Godfrey, R. M.: 'Radio Direction-Finding by the Cyclical Differential Measurement of Phase', *Journal I.E.E.*, 1947, **94**, Part IIIA, p. 705.

(2) Staff of the R.A.E.: 'Observations on the Orbit of the First Russian Earth Satellite', *Nature*, 1957, **180**, p. 937.

(3) Staff of the Mullard Radio Astronomy Observatory: 'Radio Observations of the Russian Earth Satellite', *ibid.*, p. 879.

OBSERVATIONS OF BEARING AND ANGLE OF ELEVATION OF SATELLITE I

W. C. BAIN, M.A., B.Sc., Ph.D., and R. W. MEADOWS, B.Sc.(Eng.), Associate Member.

Introduction

At the Radio Research Station, Slough, measurements were taken of the apparent bearing and elevation angle of the satellite on a number of occasions on both transmission frequencies. The object of the experiments was not so much to help in fixing the position of the satellite as to obtain information about ionospheric propagation by studying the results retrospectively in the light of exact data about the track of the satellite.

The first measurements at Slough were made on the evening of the 5th October; these were of the bearing on 20 Mc/s, and were continued at intervals until the 16th October. On the 7th October, equipment for the measurement of elevation angle on 20 Mc/s was brought into use, but unfortunately very few successful observations were made with it, largely because of difficulties with interfering stations. On the 8th October, measurements of bearing and elevation angle were begun on the frequency of 40 Mc/s at a site at Winkfield, 13 km to the south-west of Slough, and these too were continued at intervals up to the 16th October.

Equipment

Bearing on 20 Mc/s.—The bearing observations on 20 Mc/s were taken on a standard h.f. Adcock direction-finder. The receiver used was the Admiralty type FHB, which has twin channels and a cathode-ray-tube display. The reading of bearings was carried out visually and timed by a chronometer.

Angle of elevation on 20 Mc/s.—The equipment used consisted of two horizontal loops spaced vertically and connected to the respective channels of a twin-channel receiver. The angle of elevation can be derived from a measurement of the ratio of the signals from the two aerials, as described by Wilkins and Minnis.[1] There are a number of ambiguities in the interpretation of the results, but never more than four; it was found possible to resolve these without difficulty from a rough knowledge of the track.

Bearing and angle of elevation on 40 Mc/s.—Apparatus at Winkfield, which had been designed for directional measurements on long-distance v.h.f. transmissions from Gibraltar, was used at 40 Mc/s. This consisted of a pair of horizontal dipoles spaced horizontally $\frac{3}{4}\lambda$ apart in the plane normal to the great-circle path from Gibraltar, which has a bearing of 194° from Winkfield. For the measurement of elevation angle, another dipole was used, placed at a point above one of the original dipoles. From the horizontally spaced pair, signals were taken through equipment which displayed on a cathode-ray tube the phase difference between them, and from the vertically spaced pair the relative amplitude ratio was displayed on another tube.

The cathode-ray traces were photographed every 5 sec, together with a clock-face. The phase differences and amplitude ratios were read off subsequently and converted to bearing and elevation angle respectively. The principle of the elevation angle measurement was the same as that described for 20 Mc/s; further details are given in another paper.[2]

As the dipoles were contained in a vertical plane normal to the bearing 194°, the pick-up and accuracy of the system were poor within say ±10° of the bearings 284° and 104°.

Results

On 20 Mc/s.—The value of the work carried out on 20 Mc/s is

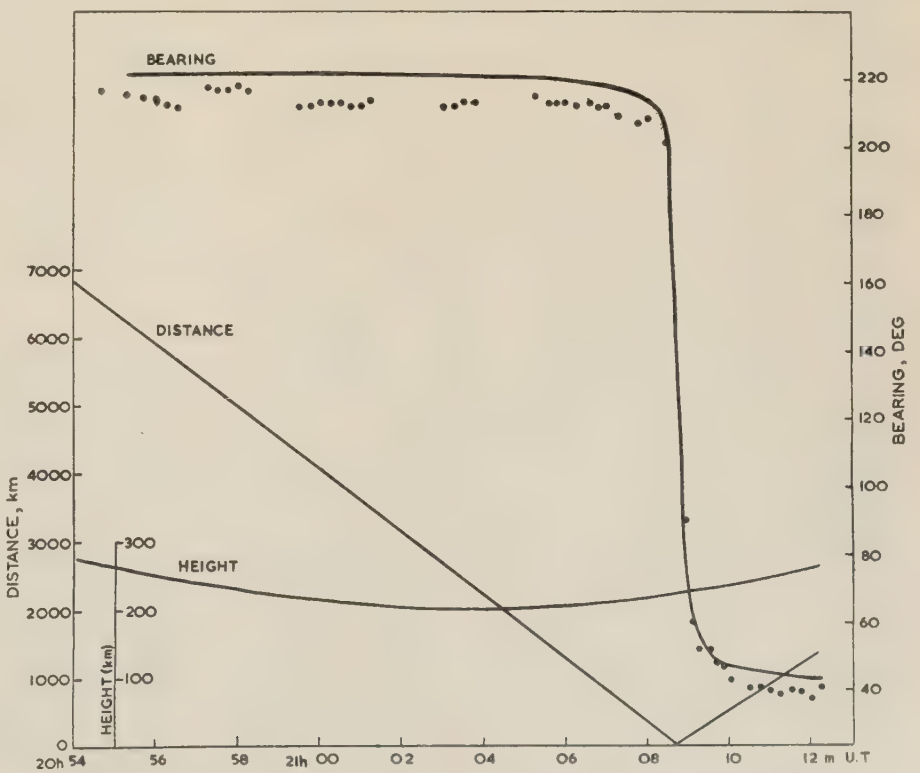


Fig. 17.—Variation with time of the bearing on 20 Mc/s, 15th October, 1957.
The height of the satellite and the distance of its nearest point of approach are also shown.

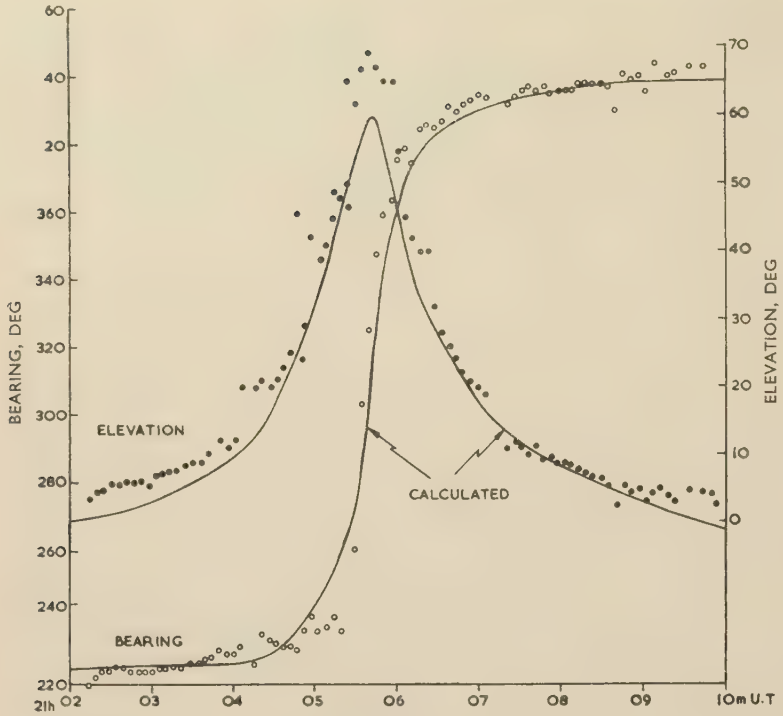


Fig. 18.—Variation of bearing and elevation with time on 40 Mc/s, 16th October, 1957.

considerably reduced by the scarcity of data on angle of elevation, which is probably of more importance than the bearing in the study of ionospheric phenomena. Accurate tracking information is not yet available here for interpreting the successful elevation runs which were obtained on the 8th October, and it cannot therefore be said whether appreciable refraction effects have occurred; it appears, however, from the data that they could not have been large.

Some interesting bearing data were obtained, particularly during the daytime on the 7th October. On one of these runs, for instance, the bearing was observed to change from 20° to 350° from 1125 to 1145 U.T., corresponding to reception of signals when the satellite was on the far side of the north pole.

The kind of comparison which is possible when the position of the satellite is known is illustrated in Fig. 17. The known curve of azimuth with time is derived from information published by the Royal Aircraft Establishment[3] and Cambridge University,[4] and bearings taken with the Adcock direction-finder are shown along with it. There are significant deviations from the curve at both ends, but these may be due in part to errors associated with the Adcock rather than to propagational effects. Incidentally, the possibility of using the satellite to help to determine errors in direction-finding apparatus should not be overlooked.

The length of time during which the signals were received is interesting. The accompanying curve for the height of the satellite in Fig. 17 shows that it corresponds roughly with the time for which the satellite was below the F2-layer maximum. The diagram also shows the distance of the satellite round the earth from the nearest point of approach to Slough; this distance is large at the beginning of the run, and reception over even greater distances has been obtained both at Slough and elsewhere. In considering these results it should be borne in mind that the geometrical horizon distance for a height of 300 km is only about 2000 km.

On 40 Mc/s.—In Fig. 18 are shown plots of experimental points giving the variation of bearing and elevation angle with time for a night-time transit on the 16th October. The solid lines represent the results of calculations based on the R.A.E. information already referred to. The agreement is surprisingly good, as the equipment was not designed for such large changes in elevation angle and bearing; for instance bearings in the region of 284° were end-on to the horizontal dipoles, which then responded only to the vertically polarized field component. This affected the accuracy of the elevation-measuring system, and since 284° was in the vicinity of the bearing corresponding to maximum elevation, the group of points at the top of the elevation/time curve in Fig. 18 is suspect.

The period during which the satellite was expected to be above the optical horizon for a given transit was calculated from the measured total change of apparent bearing for each transit, and compared with the period of time actually observed; heights of 230 km for the night and 460 km for the morning transits were assumed. In most cases the period observed was greater than that calculated, showing that diffraction around the surface of the earth, or refraction in the medium, was taking place. During the night-time runs the greatest differences occurred with the transits of lower altitude (i.e. lower maximum value of elevation angle) as might be expected, since the satellite was not then crossing the horizon as fast as it would do for overhead passages; in fact, for transits which were nearly overhead the observed period was always within half a minute of that calculated from the observed change in bearing, and for transits of low elevation the observed period of observation was as much as twice that calculated, a difference of some 3 min. These effects are probably due to diffraction and/or atmospheric refraction, since the satellite was substantially below the height of maximum ionization density of the ionosphere, the critical frequency of which being hardly high enough for oblique reflections to have occurred.

With the morning runs the same general picture was observed, except that the only near-overhead transit was received for a considerably longer time than that to be expected from the change in bearing. This may be due to ionospheric refraction, since the satellite was well above the height of maximum ionization density of the ionosphere during this period.

Conclusions

Measurements of bearing and elevation made at 20 and 40 Mc/s on Satellite I have been described briefly. Agreement between the results obtained and those calculated from orbital data was reasonable; discrepancies are, however, significant in some cases and their explanation must await further consideration of the accuracy of the radiolocation techniques used by the various research teams. Some effects attributable to the ionosphere and to diffraction have been observed; in particular, at 20 Mc/s, continuous bearings were obtained on the satellite when it traversed the north auroral region well below the line of sight, and on another occasion signals were received only when it was below the height of the F2-layer maximum. At 40 Mc/s, the satellite was usually received before it rose above the horizon and after it fell below, an effect which has been attributed to diffraction or atmospheric refraction for the night-time runs and possibly, in addition, to ionospheric refraction during the day.

Acknowledgment

We wish to acknowledge the ready collaboration of our

olleagues both in making observations and in the reduction of ne data.

The work described above was carried out as part of the rogramme of the Radio Research Board. The paper is published by permission of the Director of Radio Research of the epartment of Scientific and Industrial Research.

References

1) Wilkins, A. F., and Minnis, C. M.: 'Arrival Angle of H.F. Waves', *Wireless Engineer*, 1956, **33**, p. 47.

(2) Meadows, R. W.: 'The Direction and Amplitude of Reflections from Meteor Trails and Sporadic-E Ionization on a 1 740 km North-South Path at V.H.F.', *Proceedings I.E.E.*, Paper No. 2537 R, January, 1958 (**105** B).

(3) Staff of the Royal Aircraft Establishment, Farnborough: 'Observations on the Orbit of the First Russian Satellite', *Nature*, 1957, **180**, p. 937.

(4) Staff of the Mullard Radio Astronomy Observatory, Cambridge: 'Radio Observations of the Russian Earth Satellite', *ibid.*, p. 879.

STIMATING THE HEIGHT OF THE FIRST SATELLITE FROM RADIO INTERFEROMETER RECORDS

G. B. LONGDEN, M.A.

The radio interferometer established at Lasham has been escribed by Mr. Beresford. It was a crossed interferometer perating on the 40 Mc/s transmission from the satellite, employg pairs of aerials spaced about four wavelengths apart. The erials were sited parallel and normal to a runway which provided ready-surveyed direction of 335°. The description of the quipment and data refers to the tracking of the first earth atellite; the arrangement has since been modified.

The data are produced on a roll of paper driven at 6 in/min rough a 4-pen recorder. One channel displays 1 sec timing lses related to Greenwich Mean Time with an accuracy of illiseconds, so that the time at any point on the record may be ead to within 0·1 sec. The remaining three traces give the sine r cosine of the phase difference of the radio signals at a pair of erials.

The received signal fluctuates in strength, so that the tracking nformation which was used lies in the successive times of zeros n the traces. Path difference may be defined as the difference n path lengths from the transmitter in the satellite to two spaced erials on the ground. Successive zeros on the trace correspond o increments of half a wavelength in the path difference. For a istant transmitter such as the satellite, the ratio of the path ifference to the aerial spacing equals the direction cosine of the ngle which the sight line from the mid-point of the aerials to ne satellite makes with the axis joining the two spaced aerials. orresponding to the radio frequency of 40 Mc/s, the wavelength 24·6 ft: each pair of aerials are spaced 100 ft apart. Thus ncrements of 0·123 can be detected on the N.S. direction cosine nd of 0·061 5 on the E.W. direction cosine, since both sine nd cosine displays are available on the E.W. aerials, which ermits interleaving.

The relationship of electrical to physical length of the cables s sufficiently uncertain to leave the precise bearing at any partiular zero on the trace undetermined to a few degrees of angle. hus the records permit plotting of two direction cosines against me apart from an uncertain additive constant (see Fig. 19). he double sine-cosine display on the E.W. channel shows any eversal in the sense of change: this saves replotting entailed by ne discovery of a turning-point at a subsequent stage of the nalysis.

Graphical plotting of the data is rather slow but provides a uick method of smoothing and interpolating. Without this, imultaneous values of both direction cosines are not known. he data have been excellent and required very little smoothing. robably a more sophisticated method than graphical plotting ar extract more accuracy from the data.

The method of data reduction was devised to meet two equirements:

(*a*) To use as much of the data as possible rather than perhaps a pair of arbitrarily chosen points.

(*b*) To overcome uncertainty of lobe ambiguity and ignorance of the directional zeroes.

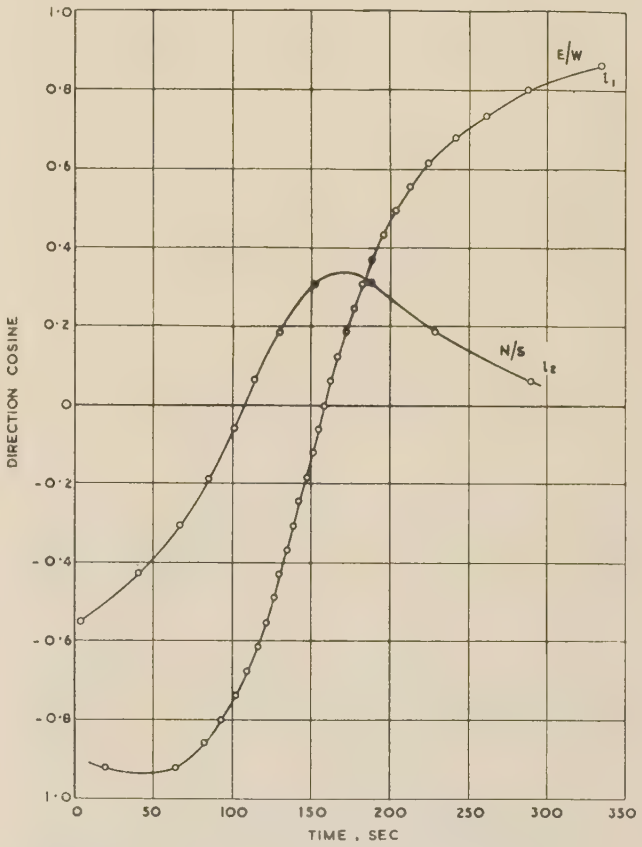


Fig. 19.—Plot of direction cosines from interferometer.

Transit SR 31: 2 054–2 100, 18.10.57.
Satellite track passed 200 n.m. S.E.

In the neighbourhood of the interferometer over perhaps 400 miles of track (100 sec) the orbit of the satellite is approximated by a constant speed v along a horizontal straight line at height h. The vector from the interferometer to the satellite may be denoted by

$$\boldsymbol{r} = r(l_1, l_2, l_3)$$

where r is the distance and l_1, l_2, l_3 are the direction cosines made with the E. axis, N. axis and vertical at the interferometer. The interferometer axes were surveyed to be perpendicular. By definition of the height h,

$$rl_3 = h$$

The projection of the vector $\boldsymbol{r}$ on the E.W. axis is rl_1 and so

the component of the satellite velocity parallel to that axis is

$$\frac{d}{dt}(rl_1) = h\frac{d}{dt}\left(\frac{l_1}{l_3}\right)$$

which is taken to be constant. Thus a graph of l_1/l_3 against time should yield a straight line.

The plots of interferometer data provide l_1 and l_2 as functions of time, provided any bending of the radio path is ignored. The third direction cosine, l_3, may be calculated from the normalizing relation

$$l_1^2 + l_2^2 + l_3^2 = 1$$

The method of analysis is to interpolate values of l_1 and l_2 from the graphs at intervals of say 25 sec and to compute the corresponding values of l_3. Then two more graphs are drawn of l_1/l_3 and l_2/l_3 as functions of time: these should be straight lines over a period of roughly 100 sec (see Fig. 20). The gradient

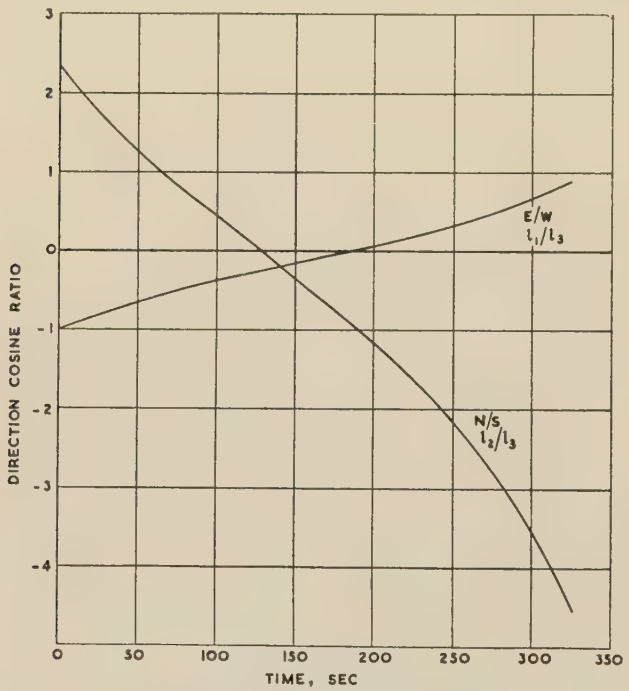


Fig. 20.—Ratio of direction cosines.
Pegasus computed values: transit at 0530, 13,10.57.
Satellite track passed 55 n.m.

of the line l_1/l_3 is the E.W. component of the satellite velocity divided by the height: similarly for the N.S. component. Since the satellite speed is a known function of height for any assumed orbit, the height of the assumed straight-line fit can be found.

Trial values for the zero levels of the two direction cosines are guessed from whatever data are available. One semi-analytical method is to assume the bearing of the satellite track and transform the two direction cosines l_1, l_2 into other direction cosines relative to axes normal and parallel to the track. These should be symmetrical and anti-symmetrical functions. The asymptotes of the one curve and the middle of the other curve can be estimated, from which the zero levels on the original plots of l_1 and l_2 can be deduced.

Errors in the assumed values of the zero levels of each direction cosine cause the plots of l_1/l_3 and l_2/l_3 to deviate from straight lines. While errors in either zero affect both curves, the greatest effect is on the ratio of the corresponding direction cosine. It has been found that an upwards concavity in the plot of direction cosine ratio can be straightened by reducing the correspondir direction cosine. With some experience, a fit of 10% accurac can be attained quickly.

Three calibration runs with an aircraft were carried out on th 15th October. Although the position of the aircraft was nc determined sufficiently accurately to calibrate the zero levels fc the direction cosines, the method of data reduction was teste for accuracy. The aircraft flew straight and level at a speed c 180 knots and a height of 15 *Kft*, roughly simulating the angula velocities encountered in tracking the satellite. The groun speed was computed from fixes over a period of 10 min to a accuracy of ½%: the height was probably correct to 1%. Th estimates of v/h for the three runs lay within 2% of the nomina values.

Together with estimates of the height, the data analysi provides estimates of the track bearing and the passing distanc i.e. the horizontal distance to the point below the neare approach. Since successive values are closely connected throug the orbital relations, a check on accuracy may be maintained b comparing successive estimates. The bearing of the track i mainly a function of the maximum latitude reached by th satellite over its complete orbit. This may be estimated fron other data which suggested bearings of 140° in the morning an 040° in the evening. The discrepancy between these bearing and the calculated bearings is a good guide to the accuracy o fitting.

The bearing of the satellite is determined by the ratio of th direction cosines, l_1/l_2. Thus the time of crossing the latitud of Lasham may be calculated. This appears preferable to th more usual meridianal crossing, since the satellite performs a integral number of complete orbits between transits.

From the near constancy of the orbit time, it was deduce that the lengths of the axes of the ellipse were changing slowly As explained elsewhere, the maximum latitude of about 64 meant that the orbit rotated very slowly in the orbital plane Thus the heights at transit were substantially constant. Withir the accuracy of 10% fitting, any trend of the observed height over the fortnight is hardly significant. The results can b regarded as successive estimates of the constant height and bearing at transit, two values relating to the morning and evening crossings when Lasham passed under different points on the orbit. From the observed orbital period and heights, it was deduced that the satellite speed relative to the earth was 4·06 n.m./s in the morning

Table 6

SATELLITE I
MORNING AND EVENING TRANSITS, 10–24TH OCTOBER, 1957

Date, October	Nominal time (hours)	Height (nautical miles)	Passing distance (nautical miles)	Track bearing (degrees)	Time of transit, latitude 51.2° N. (h.m.s.)
			Morning Transits		
11	0530	265	63	129	05.32.15
12	0530	228	31	138	
13	0530	252	25	140	05.30.16
14	0530	253	210	140	05.28.30
			Evening Transits		
10	2300	136	80	038	
14	2100	138	123 S.E.	034	21.11.20
15	2100	133	52 S.E.	034	21.08.43
16	2100	124	51 N.W.	040	21.05.25
24	1830	123	33 S.E.	039	18.41.30

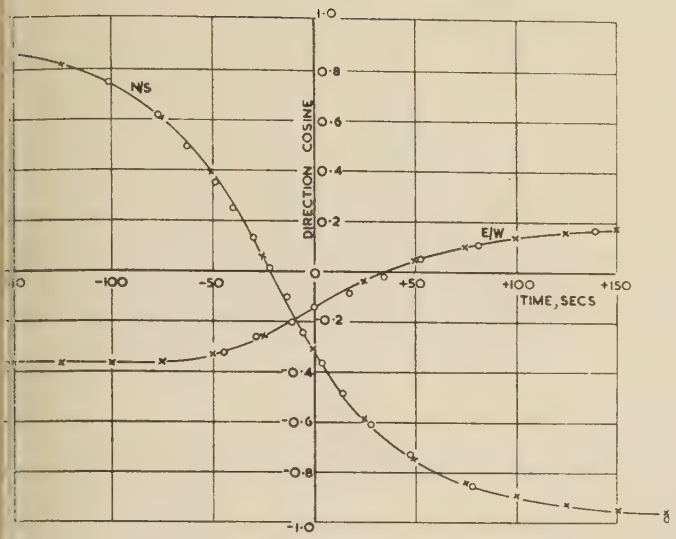


Fig. 21.—Digital computer results compared with interferometer observations.

Transit at 0530, 13.10.57.
Height, 250 n.m.
Maximum latitude, 63.7°.
Passing distance, 57 n.m.
× Computer.
○ Interferometer.

and 4·18 n.m./s in the evening. These speeds have been used in calculating the heights and hence the passing distances. The results are shown in Table 6.

For morning transits the height estimate was 250 n.m. with a root mean square scatter of 13 n.m. over four results. For evening transits, the height estimate was 131 n.m. with a root mean square scatter of 6 n.m. over five results.

The plots of the ratios of direction cosines depart from straight lines even for the best fits. An accurate calculation on the Pegasus computer based on a circular orbit over a rotating earth showed that the ratios lie on S-shaped curves (see Fig. 20). Provided attention is confined to near transits, and within about 50 sec either side of the nearest approach, a straight line is a reasonable approximation to the true curve around the point of inflexion. Several comparatively near transits within 200 miles have failed to yield any reasonable fit. The explanation may lie in undue bending of the radio path from the satellite to the ground. At longer ranges, occasional fits have been obtained, but the estimates of height are too great.

An attempt to extract greater accuracy from the data was made by programming a Pegasus computer to calculate direction cosines for an assumed set of parameters. The calculation was based on approximating the orbit by a circular arc, with earth rotation included. The trial values of the parameters were improved by visual curve-fitting and interpolation. A typical result of the fitting is shown in Fig. 21.

GENERAL DISCUSSION AT SESSION I

Mr. D. E. Hampton: From work at 20 and 40 Mc/s on Doppler phenomena it is possible to get some idea of ionospheric effects. For example, by saying that 40 Mc/s gives the right answer and that the answer at 20 Mc/s is influenced by the ionosphere, the error at 20 Mc/s may be determined. By assuming the behaviour of the ionosphere with frequency, the error at 40 Mc/s may be reduced. I have investigated the effect on interferometer and other h.f. equipment, and it seems that refraction is quite important even at 40 Mc/s. The conclusion to which I came was that the cosine of the angle of inclination of the ray was increased from the line of sight, a straight line joining the satellite to the receiving station, by about 5%. On overhead transits this is probably not very important, or at any rate on those which are immediately overhead; but if the transits are further away the height calculations can be in serious error. I shall be interested to know if anyone has done 20 Mc/s and 40 Mc/s interferometer measurements for the same transit.

Mr. E. G. C. Burt: It may be worth mentioning that we at the R.A.E. have also done some orbital calculations, and by and large they agree quite well with those which Dr. Shakeshaft has described. There are four main effects of the earth's oblateness. The first is that the recession of the nodes is $10(R/\bar{r})^{3\cdot5}\cos\alpha$ degrees per day. This is a theoretical value assuming an oblate earth, $\bar{r}$ being the harmonic mean distance of the orbit from the centre of the earth, R the earth's equatorial radius and α the inclination of the orbit to the equatorial plane. We have found very good agreement between the theoretical and experimental values, both giving a recession of about 3·18° per day.

Another effect is a rotation of the major axis of the ellipse in the orbital plane of $5(R/\bar{r})^{3\cdot5}(5\cos^2\alpha - 1)$ degrees per day. The interesting point here is that if the inclination is $\arccos(1/\sqrt{5}) = 63\cdot4°$ there is no rotation of the ellipse in its orbital plane. This is very convenient, because day after day one can measure the same height and so get a better estimate of it. For both satellites we found an inclination α of slightly less than 65°, so that there is only a very slow rotation of the major axis.

There are slight perturbations of the ellipse. The satellite oscillates twice about the mean elliptical path with an amplitude of $0\cdot94(R/\bar{r})\sin^2\alpha$ nautical miles. This is a very small oscillation, amounting to about a mile or so for Satellites I and II, and so far we have not been able to check that it is indeed there. There is an even smaller oscillation, three times per revolution, of a few feet.

Finally, the time for an inclined orbit such as this is greater by about $14\cdot5(R/\bar{r})^{1/2}\sin^2\alpha$ than for an equatorial orbit.

These theoretical results are taken from an unpublished Ministry of Supply report by D. G. King-Hele and D. M. C. Gilmore.

Mr. G. Millington: Mr. Kitchen mentioned the possibility of radiation from a satellite travelling in a leaky waveguide in the ionosphere. This suggests the use of satellites to throw light on the mechanism of the round-the-world echoes with which we are familiar in high-frequency communication.

One of their features is that each successive echo is attenuated by only a few decibels on the previous one, although the first one is usually very weak compared with the direct signal by normal F-layer reflection. They are also remarkable for their low dispersion. I have come to the conclusion that the mechanism must be that of a waveguide whereby some of the energy incident from below is trapped in the ionosphere and can travel round the twilight zone where the attenuation is everywhere low and there is a slight leakage.

When we have further satellites we must try to find out more about this type of propagation, for we have at last a means of placing a transmitter inside the ionospheric duct where this leaky mode can be excited. It is, however, only occasionally that a satellite will be in a twilight zone passing through a given point of reception and at a time when the propagation conditions are favourable.

It is interesting to note that two signals going in opposite directions round the world will have different Doppler shifts if the satellite has a component of velocity along their great-circle track. A colleague of mine claims to have observed round-the-world echoes showing this effect. I wonder whether anyone here has any evidence in support of this phenomenon.

Mr. F. A. Kitchen: I am glad that Mr. Millington took up the point about an ionospheric waveguide propagation mode. I had expected to be challenged on my remark, but he has spoken in support of it. We did have some evidence of round-the-world transmissions. On one occasion we observed a sudden transition on our bearing record, in which the signals came up on the trace 180° different from those which we had previousl received, so that it looked as though there was a transitio between signals going in opposite directions. On one furth occasion, the afternoon of the third day of transmissions fro the second satellite, we held the signal continuously on 40 Mc/ using a conventional receiver, for a number of successive orbit When the satellite went fairly nearly overhead we heard th Doppler shift, which seemed to identify the source as the satellit and show that it was not some spurious signal. This went o for two or three orbits, until we ran out of recorder tape. Thi added emphasis to the possibility of some sort of waveguid propagation mode, in view of the very low power of the transmitte

Session II

PRECISE FREQUENCY MEASUREMENTS ON FIRST RUSSIAN SATELLITE

H. STANESBY, Member.

It is easy to show that, if in free space an object emitting electromagnetic waves of frequency f_0, is travelling directly towards or away from an observer with a velocity, v, much less than that of light, Doppler effect changes the apparent frequency by an amount $\Delta f \simeq f_0 v/c$, where c is the velocity of light. For a velocity of, say, 10 km/s, the change in frequency is approximately 33 parts in 10^6. For many years astronomers have been measuring velocities of approach or recession in this way, by observing the shift of characteristic lines in the spectrum.

For an object travelling in a straight line past a stationary observer the frequency traces a skew-symmetrical curve of the form shown in Fig. 22, where the two asymptotes are spaced apart, on either side of the centre frequency, by $2fv_0/c$, i.e. by an amount proportional to the source velocity; and, referring to the diagram, the slope at the centre is inversely proportional to the distance, d_0, at the point of nearest approach. The time, t_0, at the centre is, of course, the time of nearest approach. Nowadays, frequency can readily be measured with an accuracy of 1 part in 10^8 or better, and modern techniques allow measurements to be made over short intervals which can themselves be fixed very accurately relative to G.M.T. In this brief paper it is proposed to describe the results obtained by applying these techniques to observations made on the first Russian satellite.

A satellite does not, of course, travel in a straight line. Its orbit is elliptical, its signals do not reach the earth through free space but through an ionized atmosphere, and an observer on the earth is not stationary relative to the orbit but is himself carried in a curved path owing to the rotation of the earth. However, my object is not to analyse the results in the light of these factors, but merely to describe briefly how the observations were made, and to present typical results so that others may study them.

Precise frequency measurements have been made on the satellite on 20 and 40 Mc/s, sometimes on one frequency only, sometimes on both simultaneously. Fig. 23(*a*) illustrates the equipment used for the 20 Mc/s measurements. The output of the receiver represents the difference between the incoming frequency and 20 Mc/s and is measured by counting cycles for precisely 1 sec in every 10 sec on a decimal counter. Originally this frequency was measured to ± 5 c/s and the intervals were related to G.M.T. with an accuracy of ± 1 sec. However, from noon on the 16th October these figures were improved to ± 1 c/s and ± 10 millisec or better.

Fig. 23(*b*) shows the arrangement used for the 40 Mc/s measurements. The output from a frequency-synthesizer with an accuracy of 1 c/s is preset to a value 1 kc/s removed from that

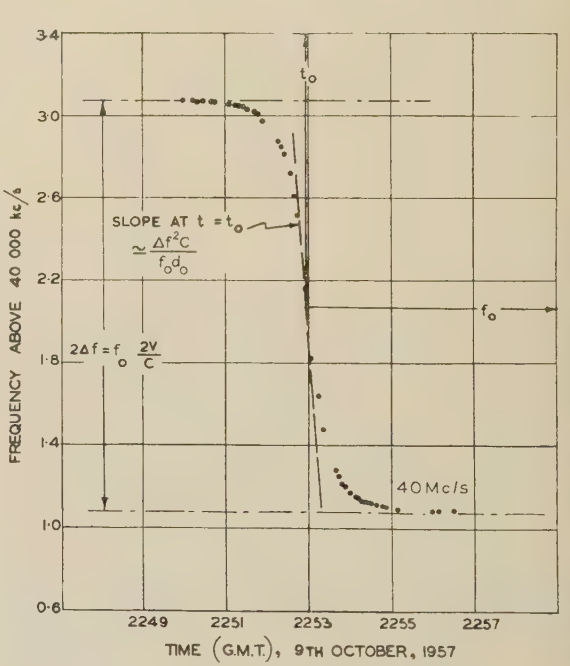
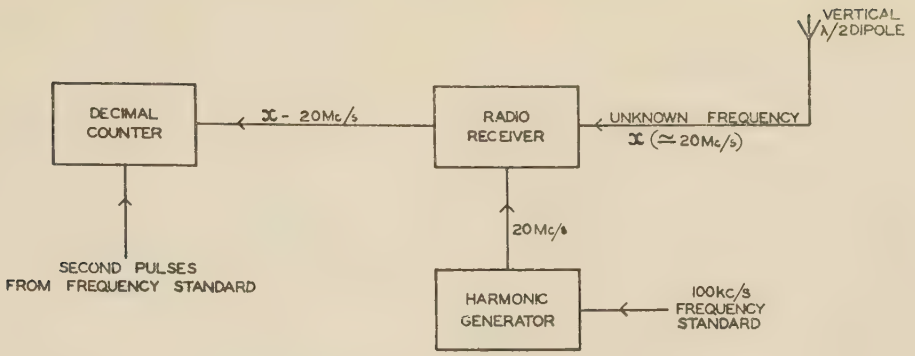


Fig. 22.—Characteristics of a Doppler curve.

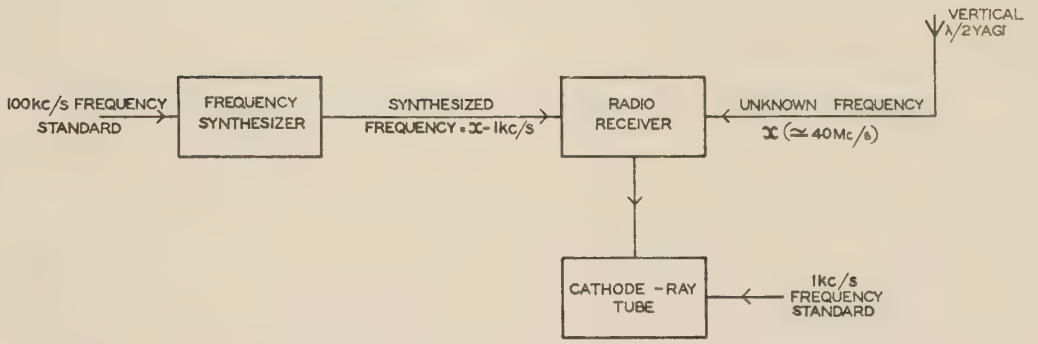
which the signal, in tracing the Doppler curve, is expected to assume a few seconds later, and the time when the beat note between the two is in synchronism with a standard 1 kc/s tone is noted from observations on a cathode-ray oscilloscope. Measurements were made at about 6 sec intervals related to G.M.T. with an accuracy of better than 1 sec.

It is, of course, to be expected that, for orbit determination, results on 40 Mc/s would be more reliable than those on 20 Mc/s because ionization will have less effect on propagation at the higher frequency. This is borne out in practice. Fig. 24 is typical of the smooth and regular curves obtained on 40 Mc/s even for a minimum distance as large as 1 700 km, although, of course, the asymptotes are not at all clear.

Fig. 25 shows the results obtained on 20 and 40 Mc/s simultaneously for a minimum distance of approximately 520 km. It is interesting to observe that the difference in the apparent times of nearest approach, $t_0(20) - t_0(40)$, is about 0·12 min and that the right-hand part of the 40 Mc/s curve shows a slight tendency to rise again.



(a) 20 Mc/s FREQUENCY-MEASURING EQUIPMENT



(b) 40 Mc/s FREQUENCY-MEASURING EQUIPMENT

Fig. 23.—Equipment for measuring frequency of satellite transmissions.

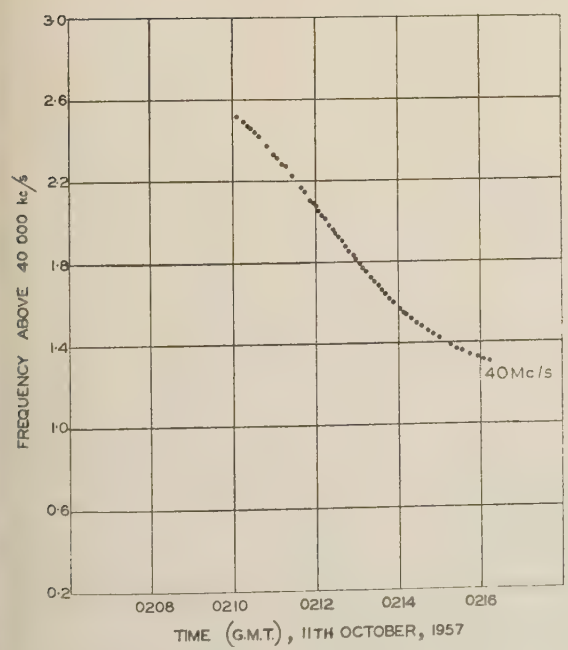


Fig. 24.—40 Mc/s Doppler curve for minimum distance of 1 700 km.

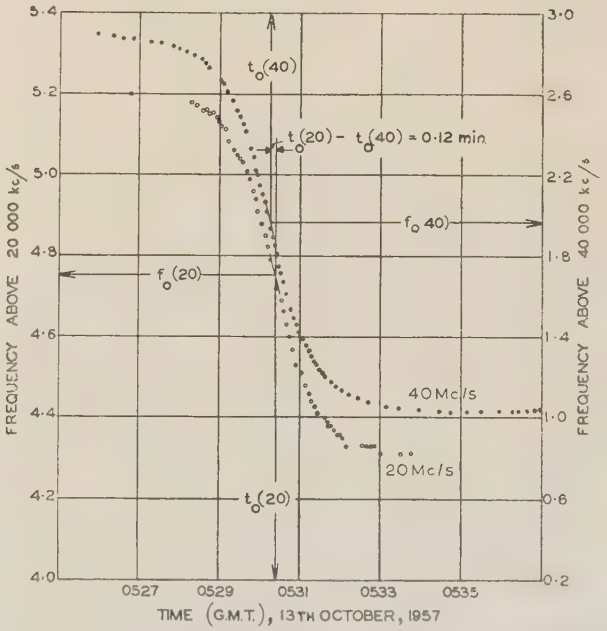


Fig. 25.—20 and 40 Mc/s Doppler curves taken simultaneously.

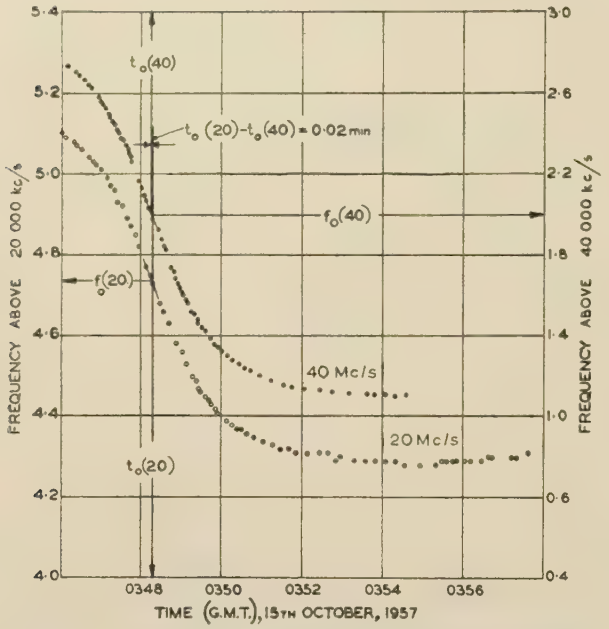


Fig. 26.—20 and 40 Mc/s Doppler curves showing small time displacement often obtained.

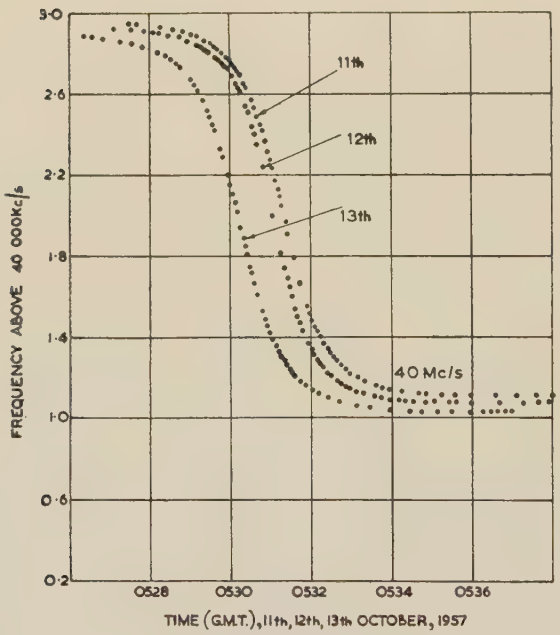


Fig. 27.—A group of 40 Mc/s curves obtained at the same time on successive days.

This tendency for the frequency to rise after the main part of the curve has been traced is exhibited more prominently in the 20 Mc/s curve of Fig. 26, which corresponds to a minimum distance of about 860 km. Here $t_0(20) - t_0(40)$ is 0·02 min, which is more characteristic of the differences in the apparent times of nearest approach usually obtained.

Fig. 27 shows a group of three 40 Mc/s curves obtained on successive days at times which happened to be very nearly the same. The minimum distances are of the order of 530 km. Such groups of curves may be of special value for orbit determination because aberrations due to propagation effects should have been similar for the three sets of observations.

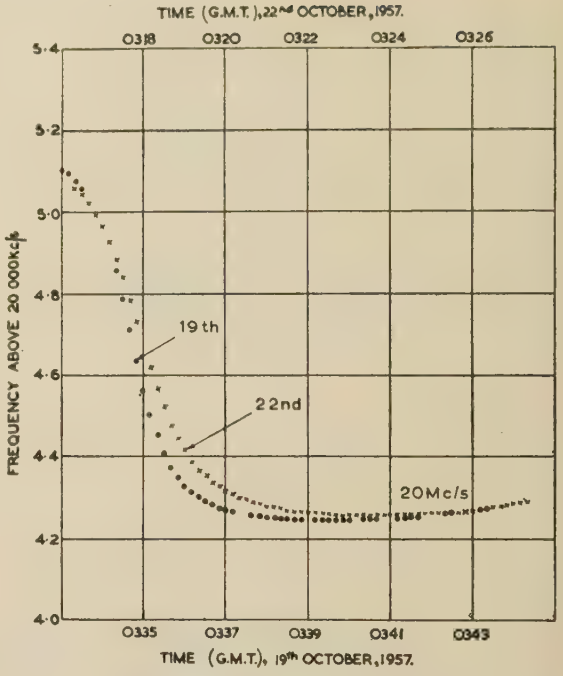


Fig. 28.—Two 20 Mc/s curves obtained at the same time on different days, showing subsequent frequency rise.

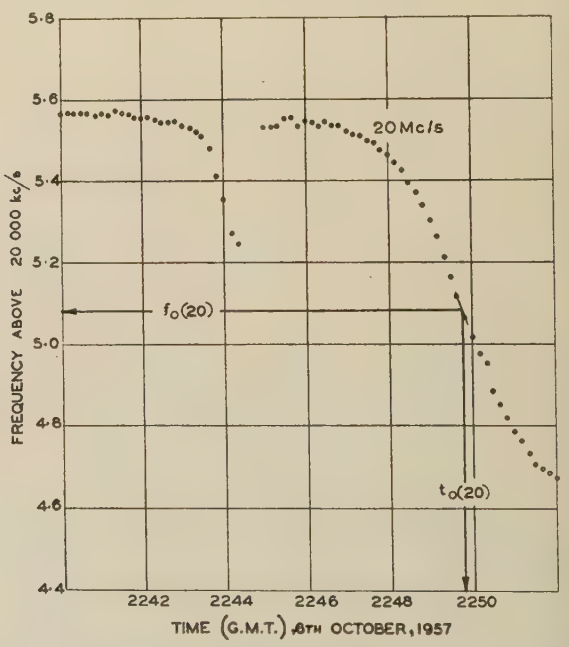


Fig. 29.—Normal Doppler curve preceded by ghost curve.

Fig. 28 shows the later portions of two 20 Mc/s curves obtained on the 19th and 22nd October and corresponding to minimum distances of about 500 and 750 km respectively. Both exhibit the rising-frequency tendency which complicates the accurate determination of the asymptotes and hence makes it difficult to determine the velocity precisely from a single set of observations. It is to be expected that, by studying the geometry involved, one might explain why the frequency rises slightly after the main part of the curve has been traced.

All the preceding material relates to measurements made at the Radio Station at Banbury. However, the last Doppler

urve, shown in Fig. 29, was obtained at our Baldock station,* n 20 Mc/s. It is particularly interesting because the main curve : preceded by the first part of a steeper, ghost, curve which ccurred about 4½ min earlier. This corresponds to a difference n satellite position of about 2000 km. I can only assume that here was then a patch of ionization near a point in the orbit bout 2000 km before the point of nearest approach to the radio tation, and that signals incident on this ionization were for a hile reflected along a path towards the station and traced a urve resembling that which would have been obtained by an bserver *in* the ionization itself. Four-and-a-half minutes later he directly propagated signals were received at the station and aced the normal curve.

Fig. 30 shows the long-term frequency variation of the 40 Mc/s atellite transmitter. There are a number of factors that may ave affected its frequency: the long-term natural ageing of the ansmitter oscillator, the periodic changes in temperature of he oscillator due to the satellite's passing from daylight to dark-ess and back again in transversing its orbit, the possibility of nock due to meteor impact, the effect of the batteries running own, and finally—an interesting possibility—the effect of enetrating radiation on a quartz crystal, which experiments show radually reduces the frequency, and by an amount that can be s much as 200 parts in 10^6.

It is interesting to observe that the points obtained for transits

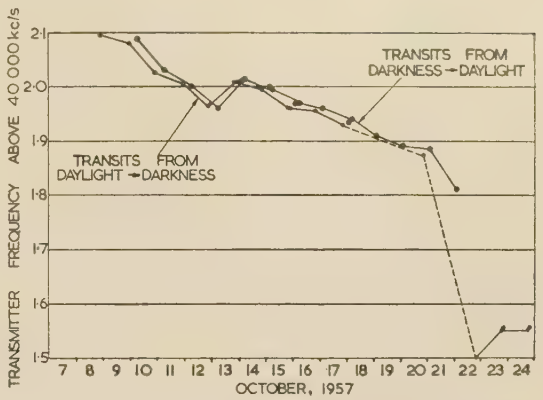


Fig. 30.—Long-term frequency variation of 40 Mc/s satellite transmissions.

following solar irradiation of the satellite and the points obtained after periods of darkness fall on different curves.

The author has pleasure in thanking the staffs of the Post Office Radio Stations at Baldock and Banbury for making the observations upon which this paper is based, and the Engineer-in-Chief of the Post Office for permission to make use of the information contained in the paper.

ANALYSIS OF DOPPLER DATA FROM EARTH SATELLITES

D. E. HAMPTON, B.Sc.

The analysis of Doppler results falls under two headings. The rst and more important is the use of the data to obtain estimates f the height and speed of the satellite. The second is the investi-ation of the effects of propagation through ionized regions. It is impossible to make a complete division between these two pproaches, since accurate estimates of the parameters for orbit etermination must involve some corrections for ionospheric ffects. However, in some cases a comparison of the results of oppler measurements on 20 and 40 Mc/s shows that quite rious ionospheric disturbances are taking place, and it is these cords which are more useful for propagation studies than for acking. The best records for tracking purposes are those ken when the transit is nearly overhead and the times of zero oppler on the 20 and 40 Mc/s records agree to within a second so. It is with the analysis of these records to obtain good esti-ates of relative speed and the minimum distance of the satellite om the receiving station that this paper is primarily concerned. y combining similar estimates on the same transit from a mber of receiving stations located in this country the height nd speed of the satellite may be estimated.

Doppler Analysis.—Fig. 31 shows a typical Doppler record. he position of zero Doppler has been found by symmetry con-derations and the frequency and time have been plotted as viations from this zero. To simplify the explanation of this rve, let ionospheric effects be neglected for the present.

The slope of the curve at the origin is determined by the la ive speed and closest approach of the satellite: the closer or ster the satellite the steeper the slope. The frequencies deter-ined by the asymptotes of the curve are related only to the la ive speed, so that by combination with the slope measure-ent an estimate may be made of the shortest distance from the ver to the satellite path. Direct measurement of slope and

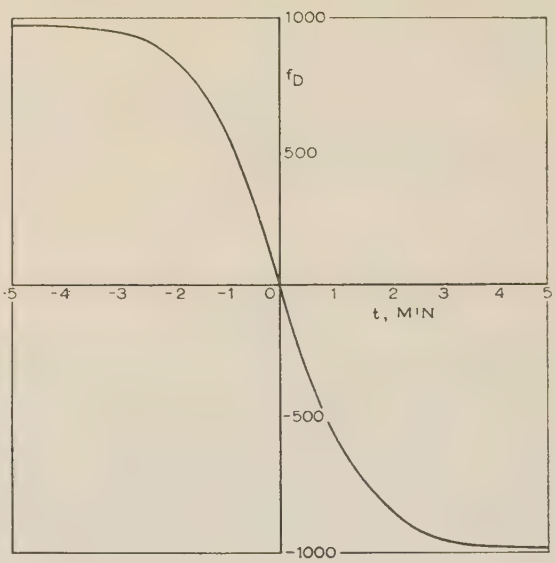


Fig. 31.—Typical Doppler curve.

asymptotes is not very accurate, and the method preferred is that shown in Fig. 32.

Here $t^2/\lambda^2 f^2$ is plotted against t^2, where λ is the free-space wavelength of the radio frequency. The results give an excellent fit to a straight line. It can easily be shown that the intercept is directly related to the slope in Fig. 31 and is equal l^2/v^4, while the slope of this line is related to the asymptote of Fig. 31 and is equal to $1/v^2$, where, neglecting ionospheric effects l is the distance of closest approach and v is the relative speed of the

* These two stations have the following positions: Banbury, 52° 02′ N., 01° 20′ W.; ck 52° 00′ N., 00° 08′ W.

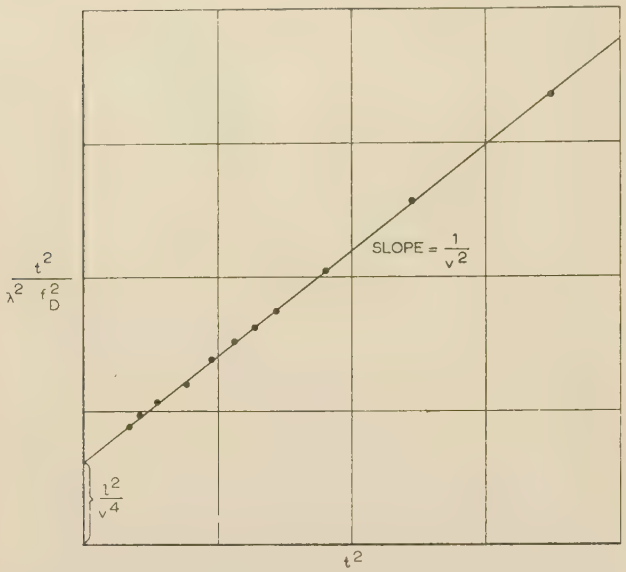


Fig. 32.—Method of data reduction.

satellite. The fit of the straight line is usually so good that there is complete justification in writing

$$\frac{t^2}{\lambda^2 f^2} = \frac{l^2 + v^2 t^2}{v^4}$$

When ionospheric effects are neglected this equation may be shown to agree in form with that obtained from an analysis of the geometry of a satellite on a circular orbit round a rotating spherical earth. However, this does not justify the assumption that ionospheric effects are negligible.

If the electrical path length l_e in a dispersive medium is defined as equal to that length of path in free space which would give the same total phase change along it, the rate of change of l_e is directly related to the Doppler frequency:

$$\frac{dl_e}{dt} = \lambda f_D = \frac{v^2 t}{(l^2 + v^2 t^2)^{1/2}}$$

Integration of this equation yields

$$l_e = (l^2 + v^2 t^2)^{1/2}$$

The integration constant is put equal to zero since an analysis of the problem gives a result similar in form to that above. Hence, from a set of Doppler measurements, it is possible to obtain the electrical path length from the satellite to the receiver.

Ionospheric Corrections.—When 20 and 40 Mc/s measurements are made on the same transit a comparison may be made of the estimates of l and v on both frequencies. Table 7 gives results

Table 7

Comparison of Results on 20 and 40 Mc/s

Date of transit	Time of transit	20 Mc/s		40 Mc/s		Corrected result	
		l	v	l	v	l_0	v_0
		nautical miles	n.m./s	nautical miles	n.m./s	nautical miles	n.m./s
10.10.57	2253	170	4·12	148	4·09	140	4·08
19.10.57	0334	274	3·96	242	3·88	234	3·85

for two particular transits, in both of which the parameters increase with decreasing frequency. Since 40 Mc/s is less subject to ionospheric effects than 20 Mc/s, the values of l from the 40 Mc/s results are the better estimates of distance of closest approach; however, they are still in error.

When the times of zero Doppler coincide on both these frequencies it seems reasonable to assume a horizontally stratified region between receiver and transmitter, so that from an analysis of propagation through such a region it is possible to show how this error behaves with frequency.

By comparing the 20 and 40 Mc/s results an estimate of this error may be made so that a better estimate of closest approach may be obtained. These corrected estimates are shown in the final column and are probably accurate to within a mile.

The deviations of electrical length l_e from true slant range arise from two principal sources: variation in phase velocity along the path, and curvature of the path due to refraction. When the refractive index in the medium is less than unity the wavelength is greater than the free-space wavelength, so that the effect is to reduce the overall phase change along the path, i.e. to reduce the electrical length. The effect of curvature of the path, on the other hand, is quite the reverse, giving an increase in the electrical path length. These results show that it is the effect of refraction which predominates, and this suggests that the angle of arrival of the ray at the receiver is appreciably different from that obtained by assuming that the ray travels along a straight path. To obtain precise orbit data from interferometer measurements there would also need to be some ionospheric correction, even at 40 Mc/s.

Conclusions.—Doppler measurements on both 20 and 40 Mc/s need to be made simultaneously to obtain accurate tracking information.

The electrical path length may be obtained from Doppler measurements.

If the ionosphere is reasonably well behaved, which can be determined from the 20 and 40 Mc/s results, a correction for the minimum slant range for the satellite may be obtained to provide accurate information for height calculations.

Acknowledgment.—Acknowledgment is made to the Chief Scientist, Ministry of Supply, for permission to publish the paper

RADIO OBSERVATIONS ON THE SIGNAL CHARACTERISTICS OF SATELLITE I

P. J. BRICE and P. N. PARKER, Associate Members.

The paper gives a brief outline, together with some comments, n field-strength observations made at the Post Office Measuring ation at Banbury, on the first Russian satellite.

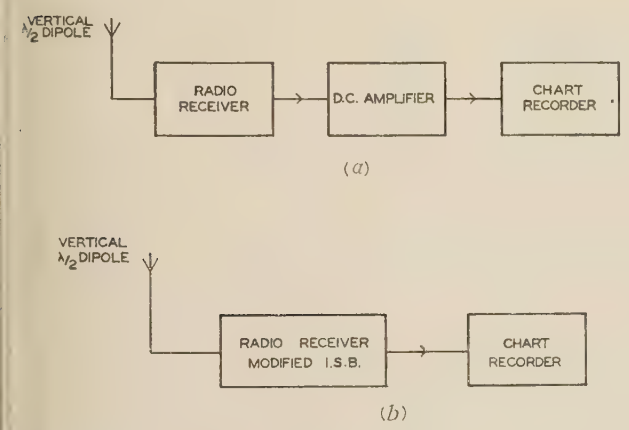
A large number of observations have been taken in the form field-strength recordings, sometimes on 40 Mc/s alone, sometes on the 20 and 40 Mc/s transmissions simultaneously. hese recordings make it possible to study the fading pattern some detail, and those reproduced here, which were taken on nkeyed transmissions, have been chosen with two points in ind:

(*a*) To show fading phenomena which appear to be of special interest.

(*b*) To correspond in some cases with simultaneous Doppler observations presented by Mr. Stanesby (see page 96). The apparent times and distances quoted here are those derived from these simultaneous Doppler observations.

The most favourable transits for observations in this country curred between 2100 and 0700 G.M.T., i.e. during night-time. this period, ionospheric absorption on 20 and 40 Mc/s is rmally very small and probably undetectable.

Measuring Equipment Used.—Fig. 33 is a block schematic of field-strength recording equipment used. Vertical half-wave



ig. 33.—Block schematics of field-strength recording equipment.
(*a*) 40 Mc/s.
(*b*) 20 Mc/s.

oles were employed on both frequencies, and the a.g.c. cuits of the two receivers were connected to a twin-pen order. A third pen was used to record 10-sec markers on same chart. The receiver sensitivity on 40 Mc/s was 1 μV d on 20 Mc/s was −25 dB relative to 1 μV.

The records shown here have been calibrated in terms of the put from the aerials into the measuring equipment, and to ive the equivalent field strengths it is necessary to consider aerial characteristics. Fig. 34 gives the computed vertical-ne polar diagram of the aerials.

t will be seen that these aerials have a poor response for y low and very-high arrival angles, and some evidence of will be seen in some of the records.

ight of Satellite During Observations.—Fig. 35 indicates the ght of the first satellite relative to the F-layer of the ionosphere the period under consideration. Data supplied by the R. indicate that the minimum virtual height, i.e. the height ich F-layer echoes from vertical soundings are first recorded, indicated by the dotted curve. The full curve indicates

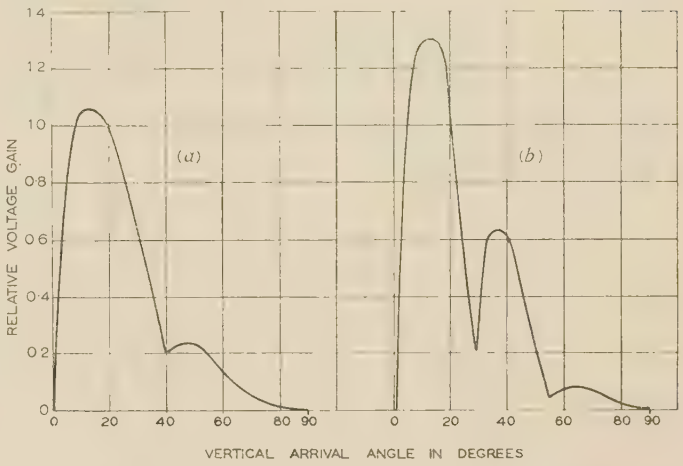


Fig. 34.—Computed vertical-plane polar diagrams of vertical dipoles.
$\sigma = 10^{-13}$ e.m.u.
(*a*) *f*, 20 Mc/s; *h*, 55 ft.
(*b*) *f*, 40 Mc/s; *h*, 37 ft.

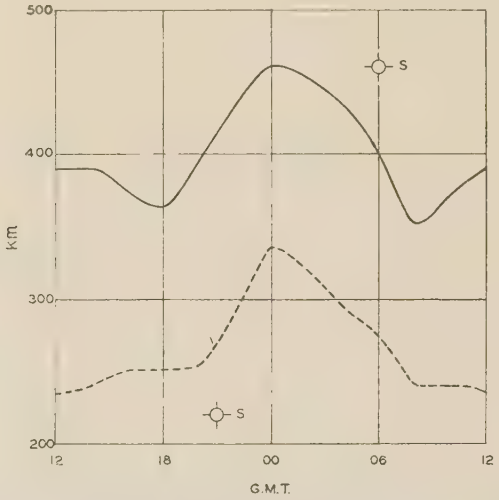


Fig. 35.—Height of satellite relative to F-layer of ionosphere, October, 1957.
——— Height of maximum ionization.
- - - - Minimum virtual height.
S. Satellite position.

the height of maximum ionization density. It will probably be helpful to remember that in the evening the satellite was at the lower height and following a SW.–NE. path, while in the morning it was at the greater height and following a NW.–SE. path.

Typical Field Strength Records.—The next eight illustrations are typical field-strength records obtained between 2100 and 0530 G.M.T., mostly on the 40 Mc/s transmissions.

Fig. 36 shows a 40 Mc/s field-strength record during an evening transit on the 10th October, when the satellite was travelling in a SW.–NE. direction at a height of about 220 km; this, of course, is not necessarily the distance of the nearest approach to Banbury. It can be seen that the form of the field-strength record is reasonably symmetrical about the time of nearest approach, some of the broad dips in the pattern corresponding roughly to

the nulls in the receiving-aerial polar diagram. After the time of nearest approach the recording shows a fading pattern with periods of about 4·5 and 9 sec. Such a pattern is also present, but to a less marked degree, on the earlier part of the record, but it is lost completely for the 1½ min prior to the time of nearest approach, corresponding roughly to a distance of 720 km. It seems beyond all reasonable doubt that the 9 and 4·5 sec periods correspond respectively to the period of rotation and half the period of rotation of the satellite. This is strongly confirmed in some of the latter recordings.

Fig. 37 shows a 40 Mc/s record for the same period the following evening. While the pattern is again symmetrical about th time of nearest approach, the fading pattern periods of approxi mately 4·5 and 9 sec persist through most of the recording.

In both Figs. 36 and 37 the period during which the satellit signals could be recorded was approximately 7½ min, corre sponding to a portion of the orbit approximately 3 600 km long Further the peak values are about 25 dB above 1 μV.

Fig. 38 is another 40 Mc/s record taken at 0530 G.M.T. o the 13th October, when the satellite was moving NW.–SE. at height of about 450 km, i.e. above the height of maximun ionization. The structure of this record tends to be markedl

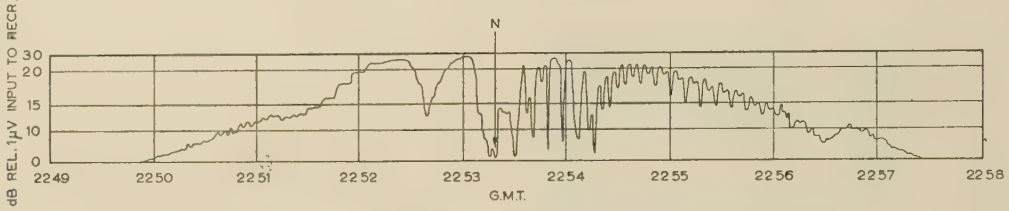


Fig. 36.—Field-strength recording: 2250 G.M.T., 10th October, 1957: 40 Mc/s.
N. Apparent time of nearest approach.

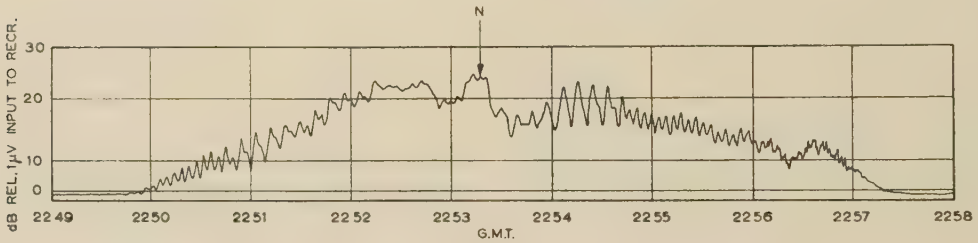


Fig. 37.—Field-strength recording: 2250 G.M.T., 11th October, 1957: 40 Mc/s.
N. Apparent time of nearest approach.

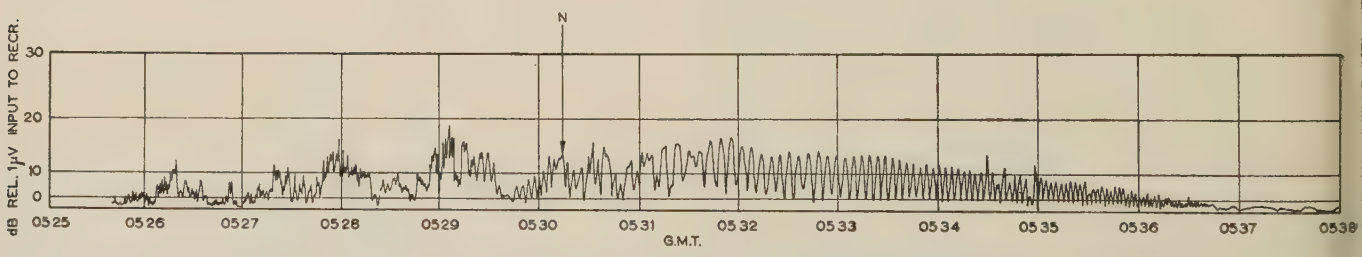


Fig. 38.—Field-strength recording: 0530 G.M.T., 13th October, 1957: 40 Mc/s.
N. Apparent time of nearest approach.

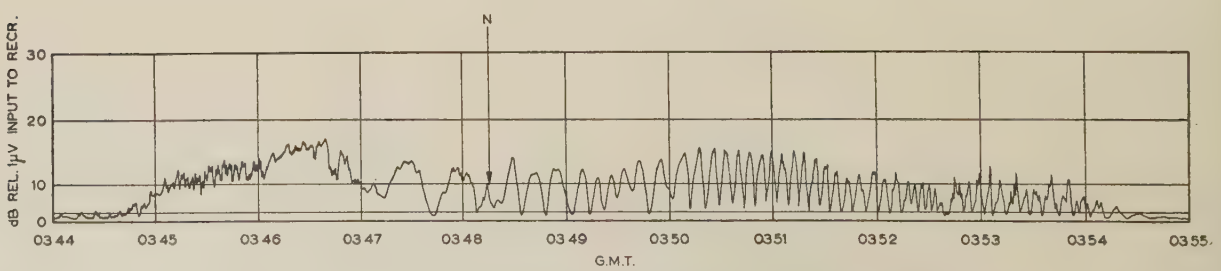


Fig. 39.—Field-strength recording: 0350 G.M.T., 15th October, 1957: 40 Mc/s.
N. Apparent time of nearest approach.

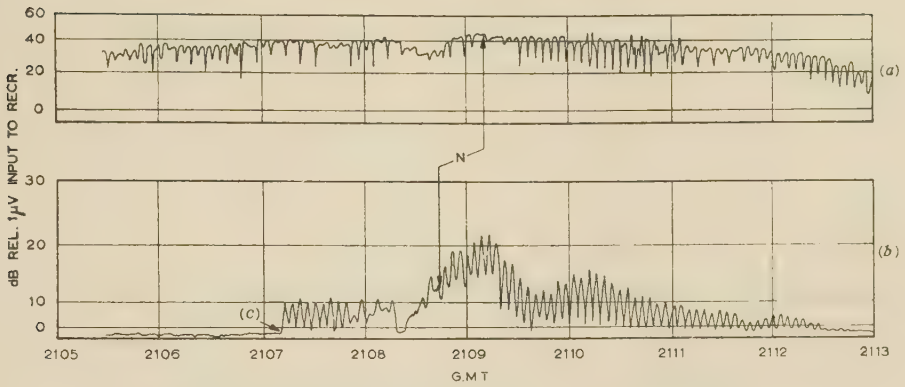


Fig. 40.—Field-strength recordings: 2110 G.M.T., 15th October, 1957.

(*a*) 20 Mc/s.
(*b*) 40 Mc/s.
(*c*) No record taken prior to this time.
N. Apparent times of nearest approach.

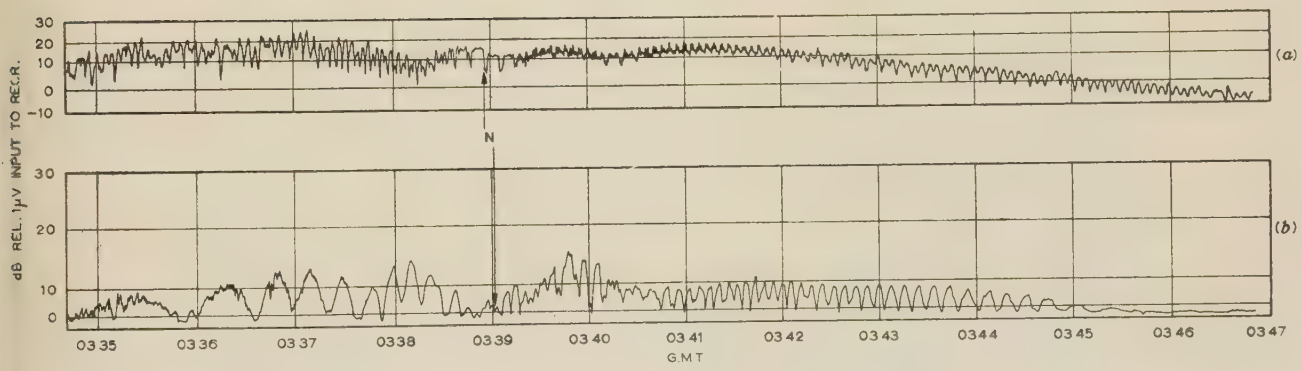


Fig. 41.—Field-strength recordings: 0339 G.M.T., 18th October, 1957.

(*a*) 20 Mc/s.
(*b*) 40 Mc/s.
N. Apparent times of nearest approach.

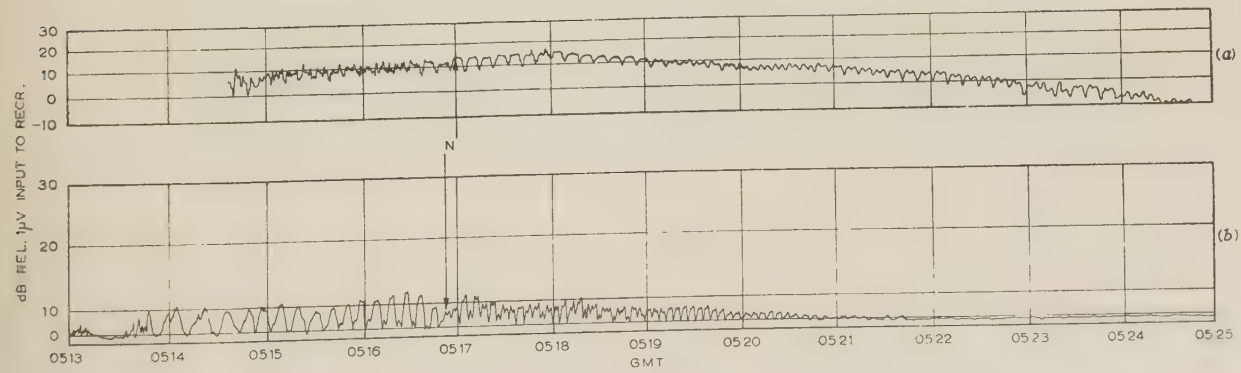


Fig. 42.—Field-strength recordings: 0520 G.M.T., 18th October, 1957.

(*a*) 20 Mc/s.
(*b*) 40 Mc/s.
N. Apparent times of nearest approach.

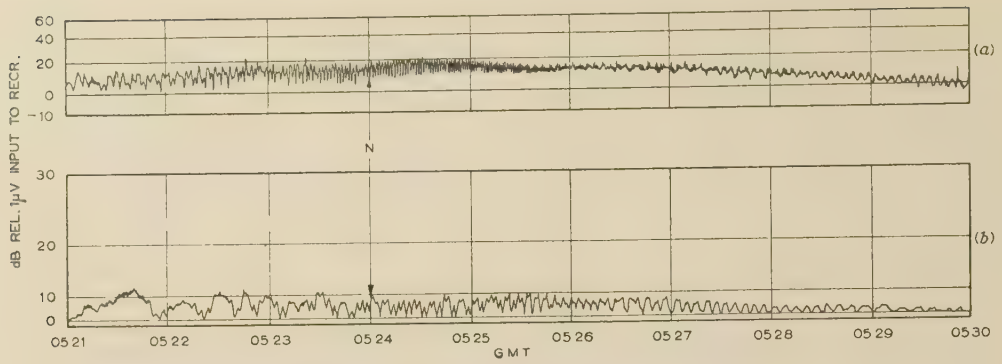


Fig. 43.—Field-strength recordings 0525 G.M.T., 16th October, 1957.

(*a*) 20 Mc/s.
(*b*) 40 Mc/s.
N. Apparent time of nearest approach.

asymmetrical about the time of nearest approach, and after 0531 G.M.T. it is very noticeable how the rate of fading increases with time.

A comparison between Figs. 37 and 38 confirms what one would expect, namely that when the satellite is higher (Fig. 38) its signal strength is lower, but that the signal can be received over a longer period.

Fig. 39 shows another record taken during the early morning of the 15th October. It is again seen that the broad structure of the record is asymmetrical about the time of nearest approach, after which the fading becomes more systematic and increases in rate.

In both Figs. 38 and 39, signals were recorded for about 10½ min, corresponding approximately to 5000 km of satellite path, with peak values of about 15 dB above 1 μV.

Fig. 40 shows recordings made simultaneously on 20 and 40 Mc/s at about 2110 G.M.T. on the 15th October, when the satellite was moving in a SW.–NE. direction at a height of about 200 km. There are two interesting features associated with these observations. First, there is a marked difference in the fading patterns on the 20 and 40 Mc/s charts, since in the former the signal strength minima are very brief, whereas in the latter the variations are more or less sinusoidal in form. Secondly, the fading of the 20 Mc/s signal has components with periods of approximately 4·5 and 9 sec, but only the 4·5 sec period appears in the 40 Mc/s recording.

Fig. 41 shows simultaneous recordings of the 20 and 40 Mc/s signals made during the 0339 G.M.T. transit on the 18th October. The fading patterns of the two transmissions differ considerably in detail. While the 20 Mc/s transmission gives a pattern in which the 4·5 and 9 sec intervals can be fairly readily distinguished, the 40 Mc/s pattern shows minima at continually changing intervals, ranging from 30 to 5 sec. A feature of the 20 Mc/s recording is a very rapid fading pattern which appears with a period of about 1 sec just after the time of nearest approach. This may possibly be due to the Faraday effect.

Fig. 42 shows the 20 and 40 Mc/s recordings for the nex transit. On these recordings both the 20 and 40 Mc/s fadin patterns, in places, show well-defined 4·5 and 9 sec periods. Th 20 Mc/s record again shows an additional pattern with a fadin period of approximately 1 sec, superimposed on the 4·5 and 9 se patterns.

Perhaps the most outstanding example of the difference in th fading on 20 and 40 Mc/s is that shown in Fig. 43, where record ings obtained at about 0525 G.M.T. on the 16th October ar reproduced. It will be seen that on 20 Mc/s, shortly before th apparent time of nearest approach, a well-pronounced fadin pattern appears, which continues with increasing frequenc until it cannot be resolved by the recording equipment. Thi seems to be one of the best examples of the Faraday effect.

Discussion of Results.—There appear to be several charac teristics that are common to a number of recordings:

(*a*) The 20 Mc/s recordings frequently show fading with com ponents having periods of 9 and 4·5 sec. The 40 Mc/s recording show somewhat similar characteristics, but in this case the shorter period component is sometimes very pronounced. These variation appear due to the rotation of the satellite.

(*b*) The recordings often show a marked asymmetry about th time of nearest approach.

(*c*) There is often a well-defined sinusoidal component of th fading pattern which changes smoothly in frequency, and withou further analysis it seems likely that this is due to the Faraday effect

Conclusions and Acknowledgment.—The records presente with this paper are representative of the large amount of field strength data obtained at Banbury Radio Station during obser vations on the first Russian satellite. Little effort has bee available to analyse and interpret these data, although comment have been made on the more obvious features.

The authors would like to thank the Engineer-in-Chief of th Post Office for permission to publish the paper. They woul also like to thank their colleagues for the generous assistanc given in preparing the material.

RADAR OBSERVATIONS OF THE RUSSIAN EARTH SATELLITES AND CARRIER ROCKET

J. DAVIS, B.Sc., J. V. EVANS, Ph.D., S. EVANS, Ph.D., J. S. GREENHOW, Ph.D., and J. E. HALL, B.Sc.

Introduction

When the first Russian earth satellite was launched on the h October, 1957, the large radio telescope at Jodrell Bank as still incomplete, and it was necessary to make emergency rrangements in order to bring it into use as a radar instrument. Assuming scattering cross-sections of the order 1–10 m^2, it as apparent that only two of the existing equipments at Jodrell ank would be capable of giving an echo at ranges greater than out 500 km. These were the 120 Mc/s long-pulse transmitter ed for lunar echo work and a small 36 Mc/s transmitter tended for use on meteor echo studies.

The telescope had previously been tested on reception by bservation of the radio star in Cassiopeia at frequencies of 38 and 408 Mc/s and the theoretical performance figures had en achieved. The performance of the telescope at 36 and 20 Mc/s could therefore be assumed to be close to the predicted gures.

Equipment

The Radio Telescope.—The aperture of the large parabolic flector,[1] which is fully steerable in azimuth and elevation, is pproximately 80 m [Fig. 44A]. For an efficiency of 60%, the llecting area is 3 000 m^2, and the aerial gain over an isotropic diator is 600 at a frequency of 36 Mc/s and 6 500 at 120 Mc/s. e full half-power beamwidths at the two frequencies are 8° d 2·5° respectively. The maximum rate of tracking is approxiately 25 deg/min; thus the instrument would be capable of llowing a satellite at ranges greater than about 1 000 km. For se of the early observations, however, the telescope was fixed a predetermined direction, and echoes were observed as the te lite passed through the aerial beam.

The 120 Mc/s Equipment.—The 120 Mc/s transmitter used for lunar observations normally operated at a peak power of 10 kW, a pulse length of 20 millisec and a pulse repetition frequency of 1 per sec (Browne *et al.*).[2] As the equipment is rather bulky, it was decided in the first instance to connect the transmitter to the aerial feed through a long transmission line, even though the overall feeder losses would be of the order of 15 dB. Under these conditions, and with a pulse length of 2 millisec, an object with a scattering cross-section of 10 m^2 would give an echo 6 dB above r.m.s. noise at a range of 1 600 km. At a later date the transmitter was mounted on the main structure of the telescope, eliminating most of the line losses and increasing this range to 2 800 km.

The 36 Mc/s Equipment.—A small 36 Mc/s transmitter giving a peak power of 10 kW at a pulse length of 10 microsec and pulse repetition frequency of 750 per sec., intended for meteor echo work, had just been finally tested prior to the launching of the first satellite. Calculations showed that, if the pulse length could be increased to 150 microsec and the feeder losses reduced to approximately 1 dB, an echo might be obtained from the carrier rocket at ranges up to about 700 km. This transmitter was mounted in the swinging laboratory at the foot of the aerial mast, directly beneath the reflecting surface of the telescope, together with the T-R switch and preamplifier. A 30 m length of coaxial cable was then sufficient to couple the transmitter to the feed at the focus of the telescope.

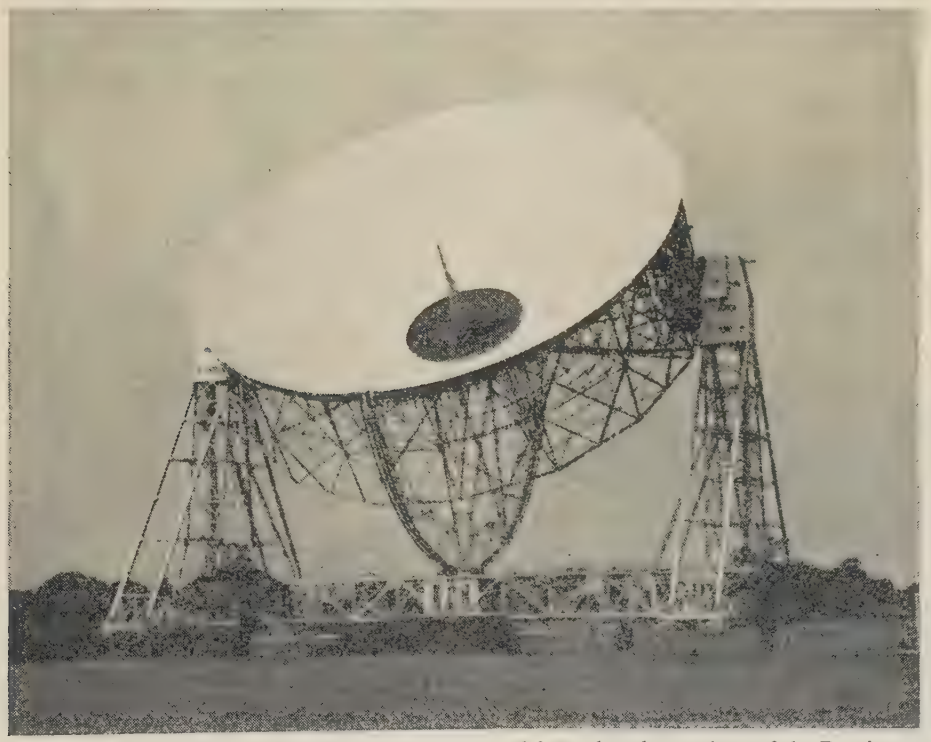


Fig. 44A.—View of the 80 m radio telescope at Jodrell Bank, used for radar observations of the Russian earth satellites.

The primary feeds for both equipments were single dipoles with parasitic reflectors, mounted with crossed polarizations to reduce the coupling between the two aerials.

The parameters of the 120 Mc/s and 36 Mc/s equipments are summarized in Table 8.

Radar Coverage of the Two Systems.—When the radio trans-

Table 8

PARAMETERS OF EQUIPMENTS USED FOR RADAR OBSERVATIONS OF THE RUSSIAN SATELLITES AND ROCKET

Frequency	Peak power	Pulse length	Pulse repetition frequency	Feeder losses	Noise factor	Bandwidth	Gain	Collecting area	Max. range fo S.-N. ~ 2 target 10 m²
Mc/s	kW	microsec	pulses/sec	dB	dB	c/s		m²	km
120	10	2000	10—20	15	5	500	6500	3000	1600
120	10	2000	10—20	5	5	500	6500	3000	2800
36	10	150	20—75	1	(sky noise) $\geqslant 13$	7000	600	3000	700

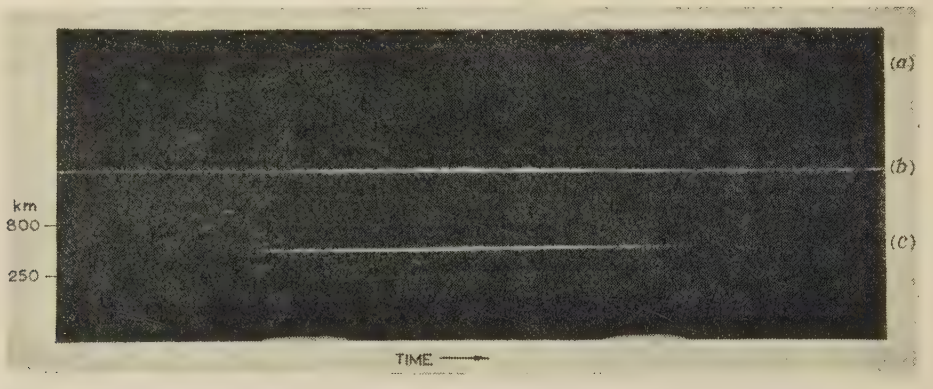


Fig. 44B.—Photographic record of a radar echo from the carrier rocket of the first satellite.
(*a*) 1 sec time marks, coded at 1 min intervals.
(*b*) Echo amplitude.
(*c*) Echo range.

mitter in a satellite is still operating it is possible to determine the orbit using radio direction-finding methods. However, in the case of the first two Russian satellites, the accuracy of the radio determinations was such that from one day to the next the time of transit of the satellite could only be predicted to within a minute or so, and its position to within approximately 100 km. When the transmitters cease operation only the inclination of the orbit and the approximate position of its plane are known. Thus, with a narrow pencil beam it is difficult to obtain a radar echo even using the predictions based on the radio direction-finding methods. In the absence of these predictions the lengthy process of scanning the plane of the orbit continuously must be adopted.

The coverage in height, Δh, and in ground range, Δl (Fig. 45) given by the beam of the radio telescope at frequencies of 36 Mc/s and 120 Mc/s, is illustrated in Table 9. It is apparent that the

Table 9

COVERAGE OF THE 80 M RADIO TELESCOPE USED AS A RADAR INSTRUMENT

Frequency	Beamwidth θ (max. to half power)	Range, R	Elevation, E	Height coverage, ΔL	Coverage in ground range, Δl
Mc/s	deg	km	deg	km	km
120	1¼	1500	10	30	75
36	4	500	30	30	70
36	4	500	90	∞	30

maximum coverage is obtained when the beam is directed at the lowest possible elevation, the limiting elevation being set by the maximum range at which an echo can be detected. Because of

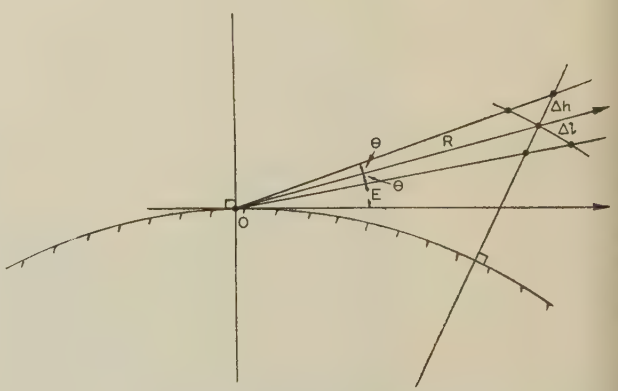


Fig. 45.—Coverage of the 80 m radio telescope used as a rada instrument at 36 and 120 Mc/s.

the wider beamwidth and lower sensitivity, the 36 Mc/s equip ment was most suitable for short-range work at high elevation while the 120 Mc/s equipment gave maximum coverage at range greater than 1000 km.

RESULTS

The 120 Mc/s transmitter was available for testing by th night of the 9th October, and the 36 Mc/s transmitter by th 10th. Some difficulty was experienced in achieving full sensi tivity at 120 Mc/s, and the first radar contacts with the rocke and satellite were made on the 11th using the 36 Mc/s equipment An example of a radar echo from the first rocket is illustrate in Fig. 44B. Echo characteristics are displayed on two cathode ray tubes, photographed by a film moving at right angles to th direction of deflection. Tube (*c*) is an intensity modulated rang display, and as the rocket passed through the fixed aerial bean

e range increased from 500 to 560 km, with a radial velocity of 3 km/s. The time of transit could be determined to within 2 sec using the 1 sec time markers also displayed on the film.

The amplitude variations of the echo are displayed on the cond cathode-ray tube, (*b*). The maximum signal/noise ratio approximately 11 dB, giving 10 m^2 as the minimum scattering oss-section of the carrier rocket.

In the echo illustrated in Fig. 44B the fluctuations in amplitude e small, although smoothing shows the presence of a shallow ctuation with a period of approximately 2 sec. Other echoes, both 36 and 120 Mc/s, show fading of variable depth with out the same period. This fading could be due to rotation of e plane of polarization of the radio wave, scintillation effects the ionosphere, or rotation of the rocket. These phenomena e being further investigated.

Up to the 22nd November, 13 radar contacts had been made th the first rocket, 1 with the first satellite, and 2 with the second satellite. Because of the confusion introduced by echoes from meteor trails, however, three of these echoes cannot definitely be confirmed. The problem of identification of the echoes will be considered in greater detail in a later paper.

It is proposed to continue observations of the satellites and rockets on a systematic basis, including a search for the first satellite, which, now that its radio transmitter has failed, is difficult to detect by other means. It is of considerable scientific importance, of course, to observe the final stages of the entry of these objects into the earth's atmosphere, and the observations will be directed with this object in view.

References

(1) Lovell, A. C. B., *Nature*, 1957, **180**, p. 60.

(2) Browne, I. C., Evans, J. V., Murray, W. A. S., and Hargreaves, J. K.: *Proceedings of the Physical Society*, B, 1956, **69**, p. 901.

OBSERVATIONS AT THE ROYAL RADAR ESTABLISHMENT

J. S. HEY, M.B.E., D.Sc.

The observations which have been made at the Royal Radar tablishment were obtained with the radio telescope which we d built during the last year or so for radio-astronomy research. e telescope is only 45 ft in diameter, which is small in comrison with modern standards, but it has the special feature at it has been designed to work at a wavelength of 10 cm. is meant that the accuracy of surface contour of the parabolic flector had to be within $\pm\frac{1}{4}$ in under all working conditions.

Following the launching of the first Russian satellite on the h October, the radio telescope was converted in about three eks into a high-power radar equipment using a wavelength 10 cm to try to get radar observations from the satellite and e rocket.

I should like to digress for a moment to refer to the special lvantage of a short wavelength such as 10 cm. The gain of an rial is $4\pi A/\lambda^2$, where A is the area of the aperture and λ the avelength. There is more power, therefore, directed on to e radar target by a factor proportional to A/λ^2. The power ux of the reflected signal is received back over the aerial aperture area A, so that there is an overall factor of A^2/λ^2 in the power the receiver. There are other factors, such as the echoing ea, σ, of the target, but assuming these are nearly constant an proved performance is obtained by having a short wavelength.

The beamwidth is then, of course, narrow, i.e. the power is ing put into a narrow beam. The beamwidth is of the order λ/D radians, where D is the diameter of the aperture. This mes out, using 10 cm over $13\frac{1}{2}$ m, to about $\frac{1}{2}°$ beamwidth. I all come back to this later, because it is very significant when ying to find a target.

There is another requirement when working at very long ranges: is necessary to have a transmitter of high power. Two factors e involved, the peak power and the pulse length, and it is the oduct of these which counts in getting long-range performance. e will assume that the bandwidth of the receiver is adjusted cording to the pulse length. In some observations low power at very long pulse length have been used, but there is an dvantage in having a short pulse length, because it is then ssible to measure the range much more accurately. In our ument, therefore, we have used a pulse length of 5 microsec d a high peak power of 2 MW.

We calculated that with this equipment we ought to be able detect the first satellite—assuming it to be a sphere of about 23 in diameter, which gives it an echoing area of something like $\frac{1}{4}$ m^2—at a range of 1 000 km. The main problem has been to know where to look, because the beamwidth is so narrow. We had hoped that predictions from the interferometer observations and from any visual observations would have attained the required accuracy to direct us on to the right path, but they were far from being good enough, and detection has depended much on chance.

I have mentioned so far only the first satellite itself. Actually the chances of detecting the rocket are better; it has a bigger echoing area, and it is possible to look at it at a longer range.

The rocket of the first satellite was first detected on the 30th October. This gave a strong echo at a range of 1 400 km. The echo as seen on the radar display is shown in Fig. 46. The marks at lower range are near ground returns appearing at the

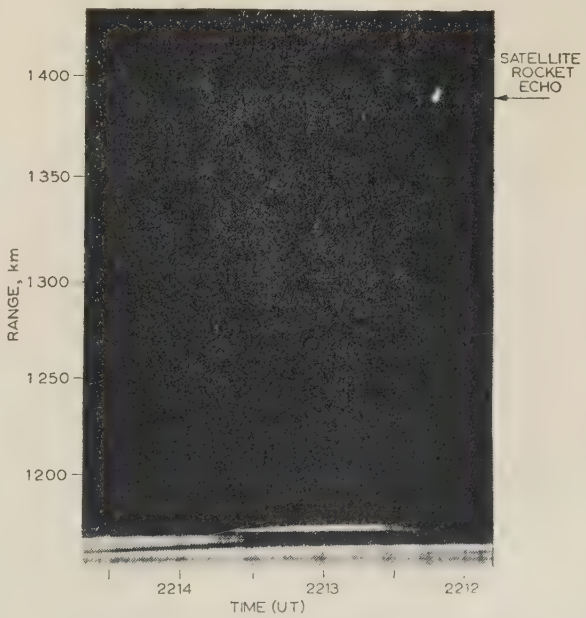


Fig. 46.—Radar echo of the rocket of Satellite I.
Obtained at the Royal Radar Establishment, Malvern, 30th October, 1957.

wrong range, because there is an overlapping of several traces of the time-base. The picture shows the appearance of the echo on the display as it moved through the $\frac{1}{2}°$ beam in the order of a second or so, and the change of range with time can be observed. It will also be noticed that it is a double echo, and this is evidence of a scintillation of echoing area due to the varying aspect of the rocket as it went through the beam.

Since then we have had a number of successful observations of this first rocket, and more recently, on the 14th November, we also obtained an echo from the second satellite. This, of course, is rather like the first rocket, because we understan that the carrier rocket is included with the satellite.

In closing, I should like to mention that not only has th Radio-Astronomy Division been concerned in this work, an particularly Mr. V. A. Hughes, but the building of this rada equipment and the receiving and recording equipment in s short a time, and the manning of the radar watches, hav involved the work of a number of people from various depar ments in the Establishment. This has, in fact, been very muc of a co-operative effort.

GENERAL DISCUSSION AT SESSION II

Mr. W. T. Blackband: I wish to comment on points which were raised in various papers about the asymmetry in time of the recorded signals from the satellite, chiefly referring to 40 Mc/s.

We have made measurements at the R.A.E. using an aerial system in which we had three dipoles arranged orthogonally—i.e. one vertical dipole, one which lay more or less along the track of the satellite, and another at right angles to it—and we made simultaneous measurements on these three. We also have observed asymmetry in signal amplitude on the separate aerials of our system. For instance, the signal duration and amplitude may be greater before than after the time of nearest approach. However, we have observed that this is not the same for the three signals of differing polarizations—in some instances a strong early signal on one aerial with weak late signal may be accompanied by a signal on one of the other aerials which is weak early and strong as the satellite recedes.

From this we conclude that any full theory of the amplitude envelope must take into account the polarization of the incoming ray, and also that measurements designed to study these effects should cover the polarization as well as the amplitude of the received signal.

There was another asymmetry which was noticed in the Post Office records, which was that sometimes there was a very deep fading on one half of the transit and very little fading on the other half, and it was noted that this was probably due to a rotation of the satellite itself, a rotation which has been observed to be about 6 cycles/min. I want to put forward an explanation of why these occurred perhaps at only one part of the transit and are at some times very much deeper than at others.

The aerial on the satellite is probably rather complex, but we can assume that it has a radiation pattern with rather large broad lobes. The satellite rotation will move these lobes with respect to an observer and produce an amplitude modulation of the received signal. This may be shallow or deep, depending upon whether it is the broad maximum or the steep side of the lobe which is directed toward the observer. If in this motion the crest of the lobe passes the observer, there will in general be two maximum signals per revolution and the modulation will be at a frequency of $2f$; if the crest is not reached the frequency will be f, the frequency of rotation of the satellite. The variation in the fades comes from the change in direction of the line of sight with respect to the satellite aerials as they pass overhead.

Mr. V. A. Hughes: At the Royal Radar Establishment, we have obtained altogether five echoes from the first rocket and one from the second satellite. Two of the rocket echoes showed scintillation whilst the others did not. The suggestion is that the scintillation is due to aspect sensitivity of the rocket. This is based on initial observations of radar echoes from the moon which have been obtained with the radar equipment. These have an extremely well-defined leading edge, indicating specula reflection, and have not yet shown scintillation that we ma attribute to the ionosphere. We hence have no real evidenc that the ionosphere produces scintillation at this wavelength.

Mr. N. Lea: I would like to refer briefly to some measure ments carried out by my colleagues at Chelmsford. The measure ments were related solely to the 40 Mc/s signal from the firs satellite. Considerations of noise and possible modulatio suggested that, after subtraction of a precision frequency o 40 Mc/s, the audio remainder might often be insufficiently clea to give good recording by direct cycle counting. A tunable l. oscillator was therefore matched to the audio remainder wit the aid of a cathode-ray-oscilloscope phase display, the outpu of the l.f. oscillator being passed to the recording systems.

Since the accuracy of the recordings depended primarily upo close following of the audio remainder by the l.f. oscillator, thre methods of tuning the latter were used, namely by a preferre slow-motion drive, by phase-lock reactance valve, and by combined manual and velodyne system. The first metho required skill but following errors could be kept within 3 c/s fo 90% of the time. The second method gave good locking bu was not used much because it was feared that phase shudder imparted to the l.f. oscillator by noise in the audio remainde would cause errors in a counter-type recorder.

The third method, illustrated by Fig. 47, required less skill an enabled following errors to be kept within 1 c/s for 90% of th time.

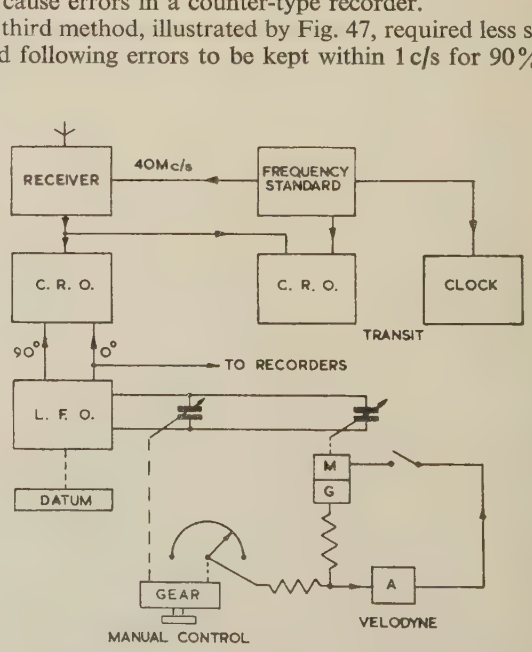


Fig. 47.—Frequency measurement plan, using combined manual an velodyne following.

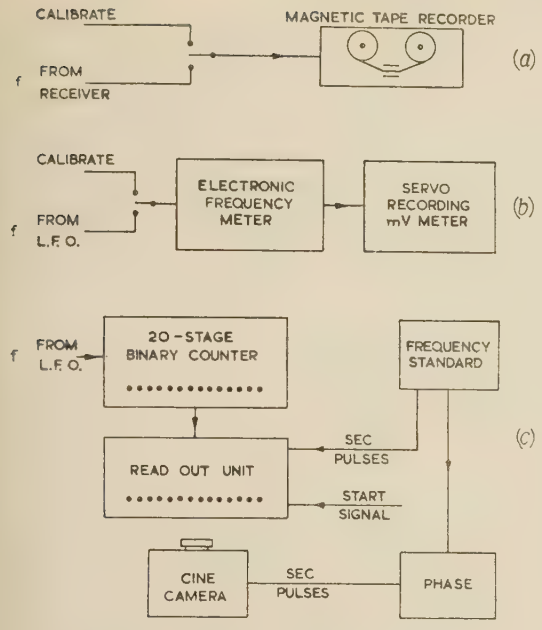


Fig. 48.—Frequency recording systems.
(*a*) Magnetic tape.
(*b*) Frequency meter and recording millivoltmeter.
(*c*) Counter and film camera.

The frequency of the l.f. oscillator was recorded automatically in two ways as indicated at (*b*) and (*c*) of Fig. 48.

The tape recorder at (*a*) is something more than a useful monitor. At (*b*) a linearized electronic frequency-meter is used in combination with an open-scale recording millivoltmeter,

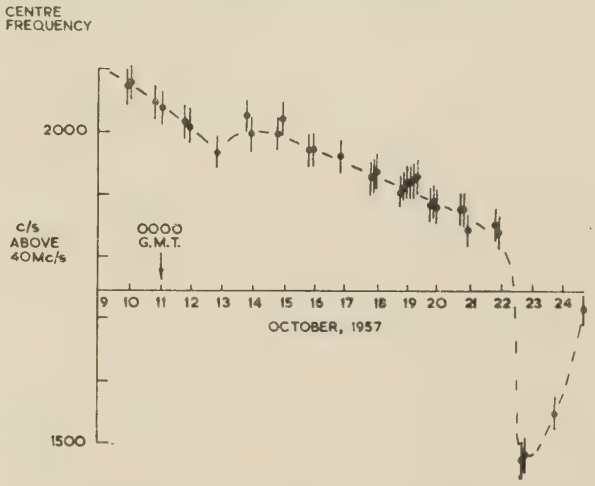


Fig. 49.—Centre frequency of first satellite.

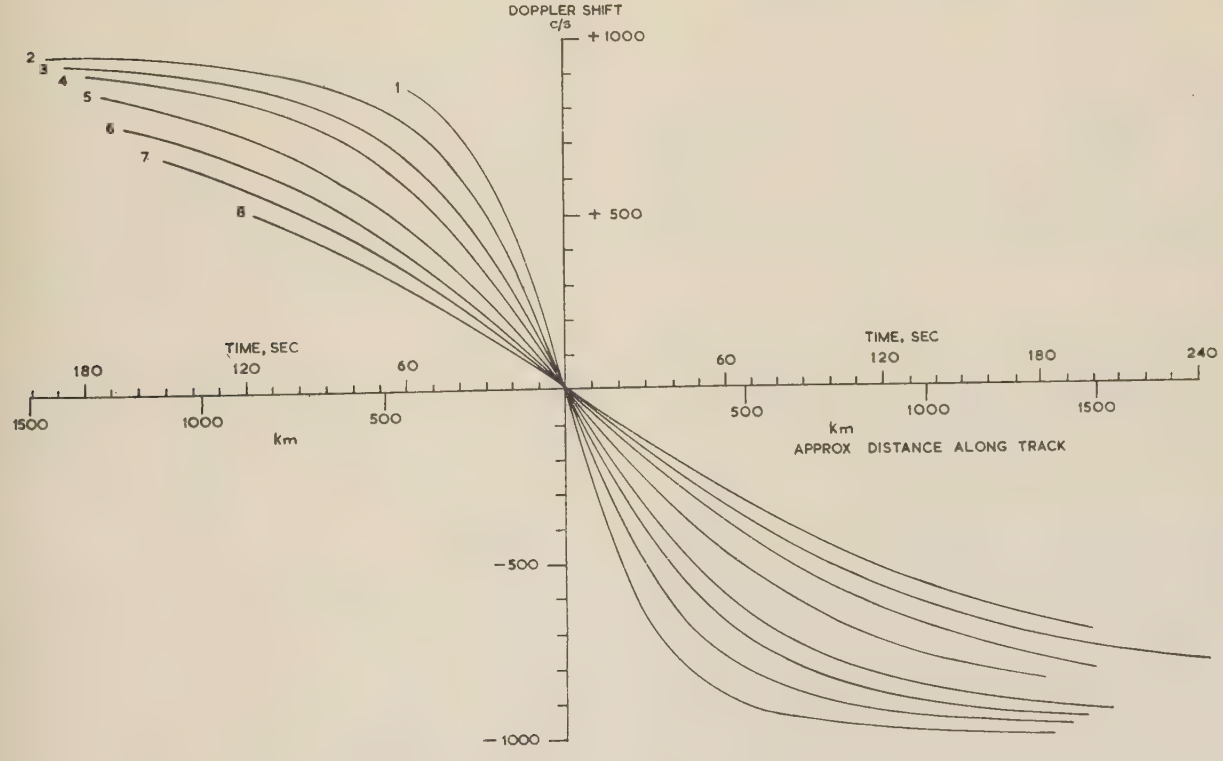


Fig. 50.—Doppler curves for first satellite, with day-to-day changes of centre frequency eliminated.

Curve	Date: October, 1957	Middle time	Slope, c/s/s
1	23	1850	28·3
2	17	2102	20·0
3	18	2058	15·2
4	19	2053	12·1
5	21	2042	9·0
6	17	1922	7·1
7	19	0156	6·1
8	19	2232	5·2

and at (*c*) the total cycles, received after a datum 'start' signal, are counted by a 20-stage binary counter fitted with a lamp-type read-out unit, the latter being photographed at precision one-second intervals. The recordings obtained by method (*b*) were remarkably clean and symmetrical.

By matching an inverted copy of a record with the original it was possible to determine graphically the centre frequency within 15 c/s, and the time of transit through centre frequency within 1 sec.

The error in estimating maximum slope appeared not to exceed 2%. Figures obtained in this way from many orbits were communicated to other laboratories.

Fig. 49 shows a plot of centre frequencies for 35 orbits, with a marked dip on the 13th October, and a much larger dip after the 22nd October, probably associated with battery failure. The vertical lines through the plotted points indicate likely limits of error in the deduced centre-frequencies.

Fig. 50 illustrates eight Doppler curves having maximum slopes between 28 c/s/s and 5 c/s/s with day-to-day variations of centre-frequency eliminated.

From the recordings by method (*c*) of Fig. 48, it is possible to make deductions of considerably higher precision. Preliminary analysis of some of these recordings by Dr. Shinn and Messrs. Blythe and Palmer suggest, for example, that the orbit of 0335 U.T. on the 19th October gave figures whose r.m.s. departure from a best-fitting quartic was 1·9 cycles. For the same orbit the maximum slope deduced was 15·06 ± 0·02 c/s/s, and the time resolution for transit through maximum slope was ±0·036 sec. It must be understood that propagation uncertainties prevent these resolutions being used directly to define accuracies of orbital information.

Mr. H. C. Woodhead: At the Post Office Measuring Station at Baldock, during the passage of Satellite I we did altogether some 200 observations on 20 Mc/s and 40 Mc/s. I should like to confirm Mr. Stanesby's remarks on the Banbury observations by saying that we also found a number of cases where there were differences between transit time as measured on 20 Mc/s and on 40 Mc/s.

Another possible explanation has occurred to me for Fig. 29, in which a 'ghost' Doppler is shown leading the main Doppler by some 5 min, or 2 000 km. Is it known whether the rocket casing was at that time circling the earth some 2 000 km in advance of the satellite? If that should prove to be the case, it might be that signals from the satellite were being reflected by the rocket casing at its nearest approach. It will be observed that the ghost Doppler gives, from its steepness, a distance of some 250 km from the source to the reflecting point, whether it occurred as Mr. Stanesby suggested or as a reflection from the rocket casing, as against some 560 km for the satellite which came 5 min later.

There were similar cases on 20 Mc/s during the early transits when the transmitter was keying on the two frequencies alternately. After the 7th October (Fig. 29 was taken on the 6th), it seemed to us that the mechanism went wrong, as if it had stuck in the 40 Mc/s position, and the signal which we got on 20 Mc/s was attenuated perhaps by the capacitance of a relay switch. I think that the figure quoted by Mr. Parker of something less than 1 watt was for this latter condition, and that originally the 20 Mc/s transmission may have had as much power as the 40 Mc/s transmission.

Mr. H. Stanesby: I am grateful to Mr. Woodhead for his remarks, confirming that the apparent times of nearest approach derived from the 20- and 40-Mc/s Doppler curves were often observed to differ at Baldock, as they were at Banbury. I am interested in his suggestion that the ghost Doppler curve shown in Fig. 29 might have been due to signals from the satellite falling on the rocket whence they were re-radiated. I do not think, however, that this could be so. Accepting that the slop of the ghost Doppler curve corresponds to a range of abou 240 km, let us suppose that the rocket was leading the satellit by 240 km, and that both were about 2 000 km away from th point of nearest approach when the ghost curve was being traced any signals re-radiated from the rocket would have been man orders weaker than those from the satellite itself. I therefor find it difficult to imagine conditions under which we shoul have received the re-radiated signals and not those directly fror the satellite.

Mr. Lea described the very interesting technique which he ha been using for making Doppler observations. I might hav stated in my paper that, when a counter was used to measure th frequency of keyed signals, an oscillator accurately synchronize with the signals by means of a highly-developed automati frequency-control system was used to bridge the interval due t keying.

Mr. H. V. Griffiths: As is well known, we were concerne with these observations from the earliest period; in fact, fror just prior to 0015 G.M.T. on the 5th October, when the satellit signals were first identified. I shall leave to my colleague Dr Phillips the discussion of some of our measurements. Th B.B.C. station at Tatsfield is equipped to measure frequency field-strength and signal direction; the latter with an oscilloscopi direction-finder. The number of frequency measurements made mainly to disclose the Doppler shift and propagation charac teristics, has been very large. In the beginning this was partl because we felt that at 0015 G.M.T. other laboratories migh not be active; we carried on thereafter hoping to learn somethin of interest ionospherically and to be of assistance to others and our early results including orbital reception times hav been acknowledged.

I would like first to afflict you once more with this recordin of the 'bleeps'. Mr. Stanesby mentioned the reverse Dopple effect, in which the frequency graph begins to rise again as th satellite recedes into the far distance, but you will first hear th signal made by a locked oscillator heterodyning the satellit transmission, as it occurs in the middle part of the Dopple characteristic. The frequency can be heard falling in pitc [Demonstration].

The next piece of the tape is interesting because it exemplifie the turn-over or reverse Doppler effect [Demonstration]. Fig. 5 shows a similar but complete transit, though not the same one The turn-over effect will be seen at the foot of the main slope.

These satellite signals have commonly been heard for lon

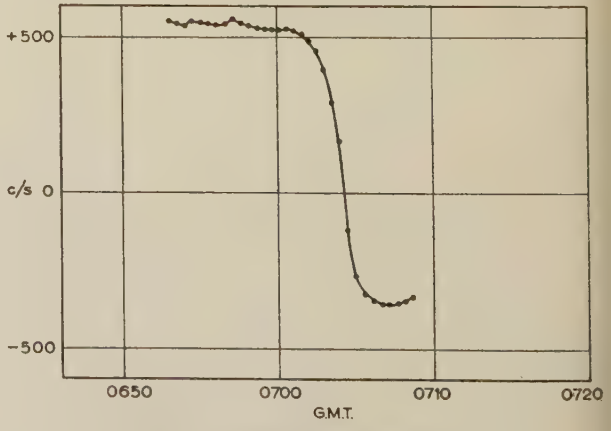


Fig. 51.—Bleeps from Satellite I.
B.B.C. Tatsfield station, 7th October, 1957.

riods continuously; as long as 40 min on 20 Mc/s and almost
5 min in one case on 40 Mc/s—of remarkable interest in
opagation study.

A further point I would like to mention is this. You are
miliar now with the regular noise of the bleeps, but there was
least one occasion (0840 G.M.T., 7th October) when the pulses,
stead of exhibiting the usual regularity with the descending
ote, had a peculiar triple rhythm or 'hiccup', which may repre-
nt some deliberate or accidental difference in transmission.
he pulse frequency and mark/space ratio normally varied with
e time of observation, as will be seen from Fig. 52, showing

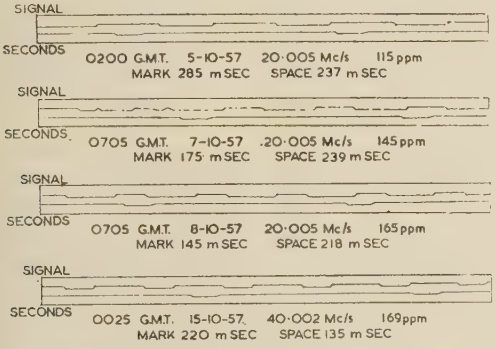


Fig. 52.—Pulse records of Satellite I.
B.B.C. Tatsfield station, 5th–15th October, 1957.

ief extracts from our pen-chart recordings. The satellite
gnals initially had a repetition frequency of 108 pulses per
inute rising later in one period to 180 pulses per minute. The
cording shows examples of different mark/space ratios and
lse repetition frequencies, with the satellite recorded on the
p pen and the one-seconds marker pips on the bottom pen:
will be seen that the upper recording shows the longest periods
mark. But, for the one definite occasion mentioned, we have
tape recording of the bleeps with the rhythmic characteristic
d this, as you will hear, sounds quite different [Demonstration].

Dr. G. J. Phillips: Fig. 53 shows that on some occasions the

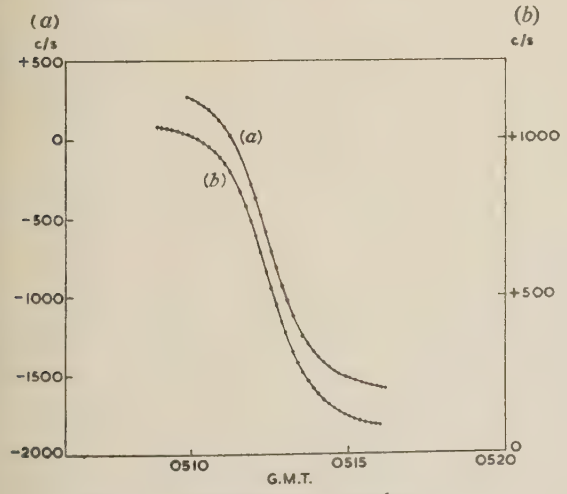


Fig. 53.—Two-frequency Doppler curves.
B.B.C. Tatsfield station, 8th November, 1957.
(*a*) Frequencies relative to 40·002 Mc/s.
(*b*) Frequencies relative to 20·005 Mc/s.

ospheric effect does not seem very large. Doppler curves
e for 20 Mc/s and 40 Mc/s are plotted with a 2/1 frequency
l, and it will be noticed that the slopes are very similar.
allowance for ionospheric refraction is made, the curves
give slant ranges of 759 km and 736 km respectively, assuming a satellite velocity of 8 km/s. Some 20-Mc/s Doppler curves plotted from the Tatsfield observations showed fairly discrete jumps in frequency. A few fit in with the idea that the satellite is just below the height of reflection, and, to stretch the term used in long-distance ionospheric propagation, we have 2½ hops and 1½ hops before the main curve, which corresponds to propagation by a direct path. Other curves are much more complicated; for example, in Fig. 54 there is an indication of

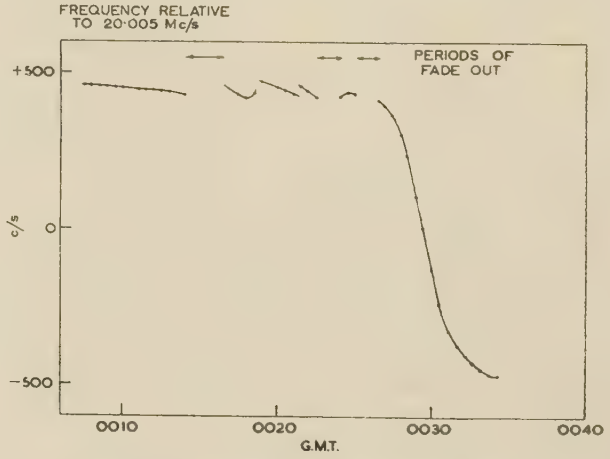


Fig. 54.—Doppler curve with several breaks.
B.B.C. Tatsfield station, 8th October, 1957.

small jumps in frequency. A possible explanation for some of these is that, if the satellite is, say, 50 km below the height of reflection, there may be a small jump due to the change-over between the signal initially being reflected from the layer above and the signal initially not being reflected.

Fig. 55 is a record of fading on 40 Mc/s taken by the Field Strength Section of the B.B.C. Research Department using two

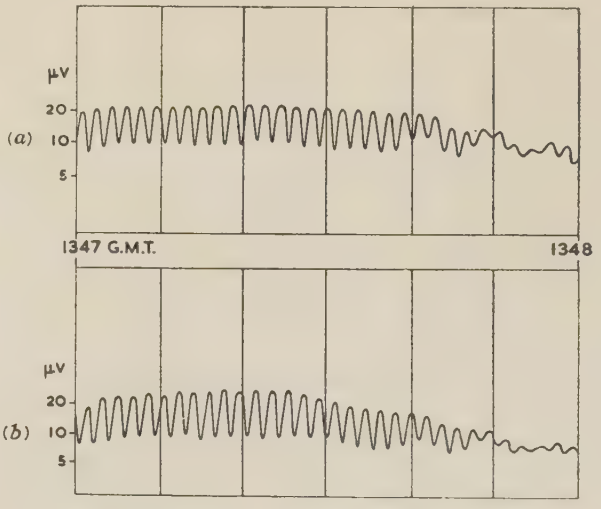


Fig. 55.—Fading on orthogonal horizontal dipoles.
B.B.C. Research Department, 6th November, 1957.
40 Mc/s voltage (open-circuit) from λ/2 dipoles, λ/4 above ground.
Orientation: (*a*) 45° from N.; (*b*) 135° from N.

horizontal aerials with orientations at right angles. Although they were in the same position, it can be seen that the fading is in antiphase on the two aerials. This would be expected for the Faraday type of fading.

In connection with this type of fading two points may be

made. The first is that diversity reception of the type which uses polarization diversity would obviously assist in any process of relaying signals. The second is that the assumption made by some previous speakers that the transmitter emitted plane-polarized waves is not in general valid. If the radiation were from a linearly polarized aerial one might expect, at any rate on 40 Mc/s, equal amplitudes of the two magneto-ionic components. The fading on a simple dipole caused by beating between these circularly polarized components would then always be deep. But the arrangement with the first satellite was that in certain directions an approximately circularly polarized wave was emitted. In those circumstances the fading would be much shallower because it would be almost purely one-component transmission. This condition may not hold, of course, for long. The polarization emitted, and hence the depth of fading, must vary with the changing aspect of the satellite. In general, an aerial for circular polarization, either at the transmitter or at the receiver, can reduce fading.

Dr. D. C. Leslie: Mr. Hampton has described how the Royal Aircraft Establishment is using Doppler measurements to study the ionosphere. I want to talk about the way in which we have used the same measurements to determine the orbit. These two activities are not independent. Mr. Hampton has to know the orbit before he can determine what the ionosphere is doing, while we have to know what the ionosphere is doing before we can determine the orbit.

The first step in the analysis is to determine the minimum passing distance from one station on one transit. Mr. Hampton has described how this is done. We find the centre frequency and then convert the data to a straight line. Because of the immense amount of data we have made a programme for the digital computer Pegasus which locates the centre by finding the minimum variance about this straight line. When we started we did not realize the magnitude of the ionospheric effect, and in fact we asked our helpers to measure only on 40 Mc/s; but this was a great mistake, because with 40 Mc/s only it is not possible to tell whether one is having trouble from the ionosphere or not. If future satellites are to have the same transmitters it will be essential to measure on both 20 and 40 Mc/s.

From the point of view of orbit determination it is no good making measurements on anything but close transits. With distant transits a small percentage error makes a great difference in height, while the line of sight is always very low and badly affected by the ionosphere. So far, therefore, as finding the orbit is concerned, there are probably only two or three transits a week which are worth looking at. I ought to say something about the accuracy at which we think we should aim. The best method is probably to set up fixed frequencies and observe the time at which the signal goes through these; a 0·1 sec in this quantity would be very good. Because of the magnitude of the ionospheric effects, we could probably not extract any more information from data which were more accurate than that.

What do we want to determine? In the first place we want the times of closest approach, so that we can give to those who are working on the orbit the information which will give them the periodic time. This can be found from the times of closest approach either on two successive transits or on two corresponding transits on following days. With either method, there is a small correction due to the two successive transits being at different distances from the receiving station. This correction can be found as soon as the orbit is known very roughly.

We also want height and the position of the track on the ground. The position information is quite independent of the time information in the procedures which we have used. We assume that α, the angle between the orbital plane and the equatorial plane, is known. It could in principle be determined from Doppler data, but the accuracy we should get would b less than that with which it is presumed to be known. Tw stations will fix the track. If one has more than two stations– on one occasion we had five—one can do a least-squares fit an get an idea of the error.

Our best results were for the 2108 transit of Satellite I on th 15th October. We made the height to be 124 $\pm$ 3 nautical mile I have two comments here. First, we do not believe that th errors are all random, i.e. they are not all observational. W think they are mainly due to the ionosphere. There is a bias i them, and the quoted standard deviation of 3 nautical miles doe not mean that we know either the error or the bias exactl The other point is that 124 nautical miles is considerably le than the interferometer determination of 131 nautical miles fc the height on the same transit. This, I think, is explained b the remarks of Mr. Hampton on the increase in the directio cosines due to the ionosphere, even though at this date th ionosphere was relatively quiet.

In doing these calculations it is not worth putting in the fu glory of elliptical orbit, rotating earth and the like: the accurac of the data does not warrant it. We make a first calculatio treating the trajectories of the satellite and the Doppler statio as straight lines, so that the curved rotating earth is replaced b a sliding plane. We then make a correction for the fact that th distance between the Doppler station and the trace of the satellit track on the earth's surface is not quite at right angles to th height vector. The only other correction of any importanc comes in because the satellite is either climbing or dropping a it passes over the observational stations. This did not matte for Satellite I, but Satellite II has a much more eccentric orbit.

I should like to acknowledge the invaluable help we hav received from the Post Office, the B.B.C., Marconi's Wireles Telegraph Co., Ltd., the R.R.E., and a number of our ow stations.

Mr. H. Gent: Dr. Hey has referred to the work at the Roy Radar Establishment on the active detection of satellites: w have also been doing passive detection work, using the 40 Mc frequency only. We have made Doppler frequency recordin and have also used interferometers, in our case a N.S. and a E.W. interferometer, both spaced three wavelengths apart an both provided with lobe-switching circuits. We have done firs order analysis on individual transits of the satellite by a relativel simple method (the main difficulty being the removal of lob ambiguities), treating the Doppler results and the interferomete results independently. Our Mathematics Division has als calculated an orbit for Satellite II which is based on Dopple and interferometer data from the R.R.E. only, and this agre as well as can be expected with the orbit published by th Nautical Almanac Office (the difference in ground track positio being about 100 km).

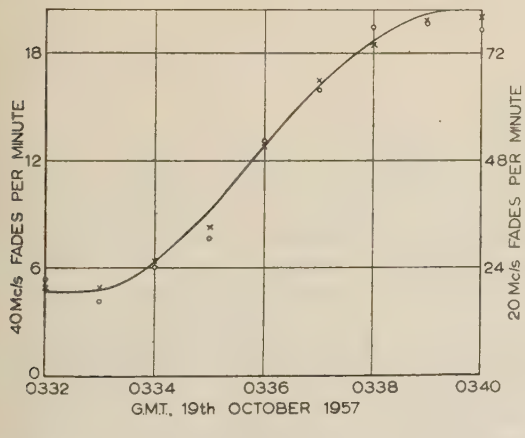
Since Mr. Hampton raised the point, I have noticed that ou Doppler slant ranges on the early morning transits of Satellite I are almost exactly 10% greater than the (independent) interfero meter slant ranges, but we have not done any investigation o this as yet.

Dr. B. Burgess: I should like to describe the analysis of som field strength measurements taken at the Royal Aircraft Establis ment on the first satellite as it passed across this country. W recorded the variations in field strength on both the 20 Mc/s an 40 Mc/s transmission during a number of evening and mornin transits. The main characteristic of these recordings, as w have heard from previous speakers, was the very rhythm fading of the received signal.

There are two possible mechanisms whereby rhythmic fadi could be produced at the receiver aerial. First, the spinning the satellite itself would cause a rotation of the plane of polariz

n of the transmitted signal with a frequency of 0, f, $2f$ r.p.m. cording to the orientation of the axis of revolution with respect the observer. Secondly, because the ionosphere is bi-refringent radio waves, the plane of polarization of the transmitted waves also rotated on passage through the ionosphere. This type of tation of the plane of polarization is called the Faraday effect, d in the atmosphere only the F-region has a sufficiently high ncentration of electrons to produce the type of fading recorded several of the transits. In order to distinguish between fading the signal due to the Faraday effect and that caused by the in of the satellite we can use the following information.

On the evening transits, the estimated height of the satellite its passage over England is about 250 km. This is some –100 km below the level of maximum ionization in the F-region, d on these transits we should have no observable fading due the Faraday effect. Thus any fading must in this case be due the spin of the satellite itself. However, on the morning nsit the estimated height of the satellite is about 450–500 km, nich is about 100 km above the level of maximum ionization the F-region. Thus the fading observed during these transits ll be due to both the Faraday effect and the spin of the satellite. us a study of the evening records will lead us to an estimate he spin. Such a study gives 0, 6·5 or 13·0 fades per minute. om this we conclude that the satellite is spininng at 6·5 r.p.m. Theoretically the rate of fading due to the Faraday effect is oportional to f^{-2}, where f is the frequency of transmission of e radio waves. Hence we would expect the Faraday effect cause fading on the 20 Mc/s transmission which was four nes as rapid as that on 40 Mc/s. Assuming this to be correct d knowing the rate of fading that can be accounted for by e spin of the satellite, we can determine the fading due to the raday effect from a study of the 20 Mc/s and 40 Mc/s record-gs of the same transit. By comparing recordings it is possible determine in any one case whether the spin had altered the e of rotation due to Faraday effect by $\pm 2f$, $\pm f$ or left it changed. Fig. 56 shows a plot of fading rate due to Faraday



g. 56.—Fading rate due to Faraday rotation plotted against time of transit.

× 20 Mc/s. ○ 40 Mc/s.

t against time of transit for the 0330 transit on the 19th tober, 1957.

We can use these rates of rotation of the plane of polarization he transmitted wave to estimate the total number of electrons t ween the observer and the satellite. We find that the mean tal electron content, in a vertical column of 1 cm^2 cross-ional area between the ground and the satellite, taken over e path of propagation through the ionosphere, is $1{\cdot}9 \times 10^{13}$ electrons. Now, radio soundings of the ionosphere at this time of day give a total electron content up to the maximum level of ionization in the F-region to be of the order of 4×10^{12} electrons. This means that we have to account for 15×10^{12} electrons in the region between the level of maximum ionization in the F-region and the satellite. This is at least three times too large a value to put in this region.

I should be very pleased to have comments on this anomaly if anyone has done any similar analysis, since this type of experiment is an exceptionally simple method of gaining knowledge of the ionosphere above the F-region maximum of electron density.

I should also like to ask if anyone has recorded any rhythmic fading of signal during the early morning transits of the second Russian satellite.

Dr. D. H. Shinn: Dr. Burgess asks if anyone has noticed that there was very little fading during the morning transits of the second satellite. Our observations at Great Baddow do in fact show this effect.

It seems to be established that the fading generally observed is due to a considerable extent to changes in the polarization of the received signal. In view of this, it is unfortunate that, as far as I know, no one has actually measured polarization. I hope that for the next satellite someone with appropriate apparatus will make such measurements.

The orbit of the first satellite is regarded as reasonably well determined, and the figures have been published. But radio measurements on the second satellite show too great disagreement with each other to enable any orbit to be deduced. However, all radio observations of which I have heard indicate a height of about 500 km for the south-going transit, 3rd–8th November. I was surprised to hear that Jodrell Bank estimate this height to be 650 km. I should like to know if anyone else has any ideas about this.

Mr. W. T. Blackband: We have not made measurements on phase directly, but comparing the measurement of field strengths on three orthogonal dipole aerials we were able to show that their relative amplitudes changed sequentially as would be expected with a rotating vector. It is an over-simplification to say that for two aerials at right angles their signals are out of phase. Their relative phase depends upon the angle which the incoming ray makes with the plane of the two dipoles.

Mr. E. D. R. Shearman: At the Radio Research Station, we have been interested particularly in the modes of propagation which enabled the signals to be received when the satellite was over the horizon. Other speakers have referred to the Doppler curves obtained under these conditions; discontinuous frequency/time curves were observed before and after the S-shaped curve representing the transit overhead, indicating that different modes were effective at different times.

An example of this behaviour was observed on 40 Mc/s at 1350 G.M.T. on the 6th November when Satellite II was travelling from north-west to south-east across England. The configuration of the track of the satellite and the F2-layer have been studied in detail for this transit, since the satellite passed nearly overhead at Slough, making the analysis of the modes of propagation particularly convenient.

The configuration for this transit is given in Fig. 57, which shows a section through the earth in the plane of the orbit and indicates the position of the observing station, Slough, the estimated height of maximum ionization of the F2-layer, and the track of the satellite. The positions of the satellite at minute intervals before and after passing over Slough are shown along the track.

The configuration shows that the satellite had climbed through the F2-layer maximum before the period of observation, and

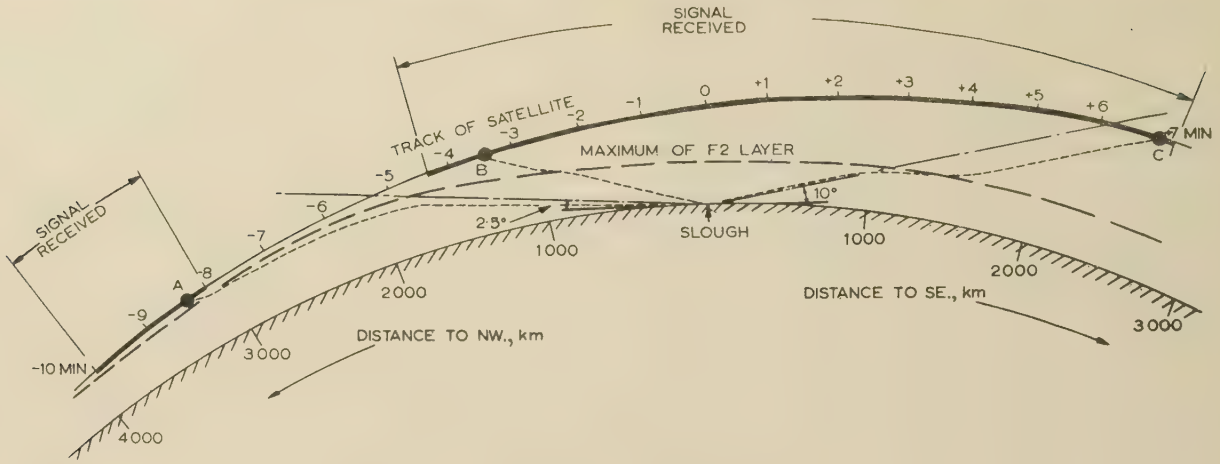


Fig. 57.—Transit of Satellite II.

Observed on 40 Mc/s at 1350 G.M.T., 6th November, 1957.

— — — — Path of radio wave from satellite to Slough. – ——— – Critical angle for escape through F2-layer.

was considerably above the layer when passing over Slough. The estimated angle of elevation for escape of a ray through the layer was 2·5° to the north-west and 10° to the south-west. Neglecting deviation of the rays, therefore, it would be expected that signals would be received when the angle subtended above the horizon by the satellite was greater than these values (trajectory as at B in the diagram). In fact the signal was first received somewhat after the satellite rose above this 'critical angle' to the north-east, but was heard for one minute after falling below the critical angle to the south-west.

The prolonged period of reception to the south-west is explicable by ray deviation of the kind shown at C. Such deviation would have been possible because the ionization in the layer was constant or increasing slightly with increase of distance in thi direction. To the north-east the ionization was decreasing fas and was therefore unlikely to sustain a deviated ray.

Before this 'line-of-sight' period of reception, signals wer also heard for two minutes when the satellite was 4 000–3 000 km from Slough. At this location the ionization was sufficientl small for a ray to penetrate the layer, undergo deviation an reach Slough by a path such as that at A. The range was to short for a '1½ hop' mode involving a ground reflection, so tha a deviated trajectory would have been necessary.

The analysis of these satellite observations has called attentio to some of the complications introduced into propagation b horizontal gradients in the ionosphere.

SUMMING-UP

R. L. SMITH-ROSE, C.B.E., D.Sc., Ph.D., Member.

If I were asked to summarize in half a dozen words what has taken place to-day, I would say that something was started seven weeks ago which has occupied the attention of nearly every radio research worker in the country, and the future of which will extend for a very great length of time.

It would be quite impossible for me to summarize what such a large number of speakers have said to-day, or to summarize the information displayed on fifty or sixty lantern slides. What I should, however, like to do is to bring out a few points about the origin of this particular activity as an international effort and see what it is leading to and where we are likely to go in the future.

First, I am certain that this particular event arose out of the planning that has been going on for the past four or five years in connection with the International Geophysical Year (I.G.Y.) —actually a period of eighteen months—which started on the 1st July, 1957. The idea of this period is that scientists all over the world, in some thousand or more observatories, shall on certain specified days of the month take measurements of the physical characteristics of the earth and its atmosphere and of the radiations from the sun which produce various phenomena here.

One part of the programme of the I.G.Y. is concerned with rockets and satellites, and so far as the British effort is concerned, the programme is confined to the release of rockets at nearly vertical incidence. Preliminary launchings have started in Australia, and later they may be followed by similar releases here.

With regard to the satellites, we may soon have to answer the question, When does a rocket become a satellite? Clearly a satellite is a body which moves in a prescribed orbit. It may or may not have been a rocket earlier because it contained some propulsive force. The I.G.Y. satellite programme was limited to the United States and the U.S.S.R. Credit must be given to the country which first demonstrated the possibility of placing a satellite in an orbit and seeing that it stayed there for, now, at least seven weeks.

This opens up a whole new field of research, as we have seen to-day, quite apart from the original objective in launching the satellites. We should not lose sight of this objective, which is to explore, in more detail than is possible from the earth's surface, the properties of the outer atmosphere; to measure the density, pressure, temperature and ionization, as well as the radiations from the sun of which we have no knowledge here, because they are absorbed in our atmosphere before they reach the earth's surface.

I do not know of anybody who has received anything which suggests what type of information is conveyed on the carrier waves with which we are so intimately concerned; but clearly

ere must be, either in the existing bodies or in some future odies, some means whereby the results of the recording instruents are conveyed to the earth, since there seems to be no rtainty at present that the instruments or their records can be covered.

In listening to the discussion, my first impression is that a ord of caution should be said about the interpretation of the sults. Many people are very experienced in measuring freuencies, and if they have the appropriate equipment they can o so to an extremely high order of accuracy. With the aid of at equipment, a number of deductions can be made about the ositions of the satellites. Other people are concerned with rectional measurements, and here again those who have had perience in measuring the direction of arrival of radio waves the horizontal or vertical plane will be aware of the difficulties d of the precautions which have to be taken before it is possible talk about fractions of a degree. Even then it is necessary distinguish between the apparent direction of arrival of the aves and the real direction of their source.

In this respect, since we are concerned with a body which is ther floating in and out of the ionosphere or rotating well ove it, we are concerned with what the ionosphere does to the dio waves which are being transmitted through it or reflected om it. It seems to me that in one or two cases investigators ve rather a naïve idea of the simplicity of the ionosphere, ich will not be shared by those who have been studying it for e past thirty years. To quote an example, some speakers seem imply that they made their measurements on 40 Mc/s because ey felt that those waves would not be affected in any way by e ionosphere; but we are at this moment at a period of maxium sunspot activity which is the highest that has ever been corded, and there can be no question whatever that we have be very careful in assuming that 40 Mc/s waves are not fected in one way or another in transmission through the nosphere. Indeed, there is positive evidence from regular observations that waves on frequencies above 40 Mc/s are being transmitted across the Atlantic at the present time, and London television signals on frequencies of this order are being received at ranges of 4 000–5 000 miles. These effects are all due to the influence of the ionosphere on the propagation of the waves and caution is necessary in dealing with these frequencies.

This applies also to the measurement of velocity. Although frequency can be measured with great accuracy, we do not know the speed of electromagnetic waves to better than about 3 parts in a million. This accuracy is for waves in free space. When it comes to transmitting them through the air under non-ionized conditions a correction has to be applied; and when it comes to transmitting them through the ionosphere, which is a doubly-refracting medium, great care is necessary in deducing a true value for the effective velocity.

What does all this amount to? First, we have been given an unusual opportunity to carry out a new type of research with the aid of a transmitter which has been provided free of all cost to this country; a transmitter which is at a greater height than we have ever had before. We should be careful to use this opportunity to the best advantage; but, when we make our radio measurements, we want to know where the source is. It would be preferable to have this position determined for us by a non-radio means, some means which astronomers should be able to provide in which they are not concerned with the troubles of the transmission of radio waves through the atmosphere and ionosphere. But when we ask the astronomers, 'Where is the rocket?' they reply, 'We do not know until you tell us where to look.' They are far worse off than the people with the radar telescopes in that their angle of resolution is very much smaller. The point is this, that radio observation can be used in the first place as a putter-on for optical telescopes, and these should then be used to determine the orbit. We should then try to relate the radio observations with the known position of the satellite in its orbit.

THE SUPPLY AND TRAINING OF TEACHERS FOR TECHNICAL COLLEGES

The Willis Jackson Report

The Report of a Special Committee appointed by the Minister of Education was discussed at an Ordinary Meeting of THE INSTITUTION, *7th Novembe* 1957. *An account of the discussion was published in the February issue of Part A of the* PROCEEDINGS. *Reprints of this account, together wi the matter which appears below, will shortly be available, price 1s. 3d. (post free). Applications for these reprints, but not for the Report itse should be addressed to the Secretary of The Institution. The Report,* price *4s., can be obtained from H.M. Stationery Office.*

THE COMMITTEE

(*a*) Willis Jackson, D.Sc., D.Phil., Dr.Sc.Tech., M.I.Mech.E., Vice-President I.E.E., F.R.S. *Chairman.*
(*a*) (*b*) R. McKinnon Wood, O.B.E., M.A., F.R.Ae.S. *Vice-Chairman.*
(*a*) Frank Briers, B.Sc., D.Phil., F.R.I.C.
(*a*) Sir Hugh Chance, M.A.
(*b*) J. G. Docherty, D.Sc., A.M.I.C.E., M.I.Mech.E.
(*b*) A. E. Evans, O.B.E., B.Sc., A.Inst.P.
(*a*) Alderman G. B. Jones, M.B.E., A.M.I.C.E., F.R.I.C., J.P.
(*b*) H. Wyn Jones, M.A., B.Sc.
(*a*) D. R. Mackintosh, O.B.E., B.Sc.
A. MacLennan, B.Sc.
(*b*) Prof. R. A. C. Oliver, M.A., B.Ed., Ph.D.
(*a*) Miss Anne Shaw, C.B.E., M.A., M.I.Prod.E.
(*b*) E. W. Woodhead, M.A.

(*a*) Member of National Advisory Council on Education for Industry and Commerce.
(*b*) Member of National Advisory Council for Training and Supply of Teachers.

SUMMARY OF THE REPORT

The Committee was constituted in the autumn of 1956 with the following terms of reference:

To consider in the light of the White Paper on Technical Education (published in February, 1956) the supply and training of full-time and part-time teachers for technical colleges, and to make recommendations.

The main objectives of the White Paper may be summarized as:

(*a*) An increase in the annual output of students from advanced courses, visualized mainly as full-time and sandwich courses, from the 1955–56 figure of 9 500 to 15 000.
(*b*) An increase in the number of students attending part-time day courses during working hours from the 1955–56 figure of 355 000 to 700 000.

This programme of expansion is to be associated with a building plan for the technical colleges of which the last projects are to be commenced by 1960–61, and of which the cost, inclusive of laboratory equipment, is expected to be of the order of £100 million.

Expressed in terms of teaching work—student-hours—this represents an increase of roughly 50% over that of the 1955–56 session.

The Committee ascertained that in the latter session the colleges had about 11 000 full-time teachers, not including the 391 principals and vice-principals, and some 40 000 part-time teachers. The number of full-time teachers had increased at an average rate of 770 per annum, and that of part-time teachers at about 2 000 per annum, during the preceding three years. But although these had been encouraging rates of growth, they had not been adequate to deal effectively with the increase in student numbers, and a mere continuance of them would fall far short of satisfying the requirements of the situation anticipated in 1960–61.

The estimates of need up to this date to which the Committee were led amount to 7 000 additional full-time teachers of a grades, and 8 000 additional part-time teachers, above th 1955–56 figures. The percentage increase in the number of ful time teachers—some 64%—is much greater than that of part-tim teachers—20%—in consequence of the changing character of th colleges' work in the direction of more full-time and sandwich course work, and of an extension of part-time day as again evening study, both of which—in the opinions of the college and the committee—can be handled most effectively by full-tim staff, if this can be made available. Converted into annua figures, this means that the number of full-time teachers *in servic* will need to increase over the five years in question at the rat of 1 400 per year, as compared with the recent annual rate of 77(A maintenance of the 1955–56 ratio between the contributions o full-time and part-time staff, on the other hand, would reduc the need for additional serving full-time teachers to 5 500 an the annual rate of increase required to 1 100.

The question which follows from these data is: What do the mean in terms of actual *recruitment* when account is taken of th likely rates of death and retirement, and of wastage from th colleges to other kinds of employment? The evidence availabl suggested an annual loss rate of 6% of the number of servin teachers, which, when applied to the increasing teacher strengt envisaged, means an average annual recruitment of full-tim teachers over the 5-year period of 2 300 or 1 800 on the tw criteria stated above. These figures are to be compared with a recruitment rate of 1 300 during recent years.

It was evident to the Committee that greatest difficulty wa likely to be experienced in obtaining an adequate recruitment o well-qualified teachers of science and technology for professiona courses at the intermediate and advanced levels. It therefore made a special study of this problem, and deduced that the recruitment of mathematics and science graduates will need t rise from the recent annual rate of 300 to a figure of 400 per yea and of technology graduates from 150 to 360. Moreover, although outside the purview of the report, note must be taken of the considerable university expansion now in progress which seems likely to call for the recruitment of some 200 graduate scientists and 150 graduate technologists per annum during the next few years.

Although the recent technical-college recruitment figure for mathematics and science graduates of 300 per year has been a gratifying achievement from the point of view of the colleges themselves, the fact that it has occurred mainly through transfers from the schools gives cause for serious anxiety. So far as technology graduates are concerned, the increased number required can come only from industry and appropriate Government Departments, since practical training and experience are essential for teachers of technology.

It is evident that, for some time to come, industry must be willing to accept, and indeed to encourage and assist, the transfer to full-time teaching work of experienced staff members it can ill afford to lose, as the only means of ensuring a much augmented future supply of junior recruits of high quality.

The Committee made no recommendations on salary scales, nce with the recent changes made by the Burnham Committee, is open to local authorities—by using to the full their dis- etion in the grading of posts and in the awarding of increments a first appointment for previous experience—to pay salaries to ll-time teachers which compare more favourably than is nerally supposed with those available in industry. But it found at transfer from industrial employment to full-time teaching ork is often hampered by pension considerations, and the ommittee hoped that arrangements might be made for a aid-up policy to be granted to those who resign to take up aching, so that a pension would be due to them on their entual retirement.

A considerable section of the report is devoted to a number matters concerning the conditions of service in technical lleges on which the Committee feel that effective action by the lucation authorities and the colleges themselves is required if rvice in technical education is to be made adequately attractive men of high qualification and ability. These matters include: e avoidance of excessive teaching loads; the more generous rovision of clerical, workshop and laboratory assistants; nproved opportunities for conducting research work, desirably close collaboration with local industry, for periodic return to dustry for further experience and for the carrying out of munerative consultative work.

The contribution of part-time teachers, particularly those in dustrial, commercial and Government employment, will con- nue to be of great importance in helping to maintain an essential ose contact between the colleges and the outside world, and to sure that the teaching takes proper account of day-to-day dustrial and commercial experience and of the latest develop- ents. With the rapid growth of day courses, their help during e day-time, and not merely during the evening, will become creasingly necessary.

The extent of this need will be affected, of course, by the ability the colleges to recruit full-time staff, but the need will in any ent be large, and the report emphasizes that, unless employers e prepared to arrange on a greatly increased scale for suitable embers of their staffs to teach part-time during the day, the lleges may find themselves unable to make the educational rovisions required. As a particular aspect of this help, the ommittee has recommended that carefully selected senior staff embers of industry, commerce and Government Departments ould be brought into close and responsible association with e academic activities of the colleges of advanced technology nd the regional colleges through part-time appointments carrying special status and title, as is common practice on the Continent nd in Russia. It visualizes that they would help, not only with e advanced teaching work, but also with the formulation of ademic policy; that they might carry a title such as Special r Associate Lecturer or Reader and receive an appropriate onorarium.

The report emphasizes the importance of a high quality of aching, and discusses the desirability and practicability of a rge-scale expansion of the facilities for technical-teacher aining. While recognizing that not all technical teachers need, or would be willing, to undergo training, and that it would not be practicable in any case to insist on this under present circumstances, the Committee feels that a greater number, particularly of those to be engaged in teaching at the less advanced levels, should be encouraged, and be afforded more generous financial incentives, to undertake it.

The position in session 1955–56 was that, of the 11 000 full-time technical teachers, about one-third had received teacher training, but of these the great majority were teachers of subjects other than technology who had been trained for service in schools, either in the university departments of education or the teacher training colleges.

The facilities available for technical-teacher training specifically comprise colleges in Bolton and Huddersfield and the Garnett College in London, which together accommodate about 300 students. An extension of these facilities, involving new buildings and residential accommodation to a capacity of 500, is recommended, this figure to be reviewed in two years' time, when it is likely that a fourth college should then be established in the Midlands.

The present course is pre-service and of a year's duration, and it is suggested that, in supplementation of this, there should be introduced for teachers already in service a 3-term course, conceived as a whole but capable of being taken one term at a time at intervals of a year or more. It is recommended that on secondment for participation in the latter course, teachers should continue to receive full salary plus an allowance towards the additional cost of maintenance. The improved grants recommended for pre-service training are designed to reduce the disparity in financial circumstances which would otherwise exist between those who wish to train before beginning to teach and those who go direct into teaching posts and take in-service training at a later date.

An associated proposal to the Minister is that a residential staff college should be established in which senior staff members of the colleges and representatives of industry, commerce, Government Departments and the universities and schools could live together for periods of a few weeks to discuss the aims, needs and methods of technical education in the light of experience and developments in these related fields.

In view of the urgency attaching to publication of the report, the Committee was unable to give as much attention as it wished to many matters falling within its terms of reference, and it therefore recommends the setting up of a permanent advisory committee so that these matters can be pursued and the whole subject of technical teachers kept under continuous review. The task disclosed by the inquiry is a difficult one, but one which must be accomplished. Its accomplishment will require the forging of a much closer partnership between the technical colleges and industry than has yet been achieved, and in the opinion of the Committee: 'There must come to be a much greater interchange of staff and ideas between them, and it must become a commonplace for individuals to move from one domain to the other, and back again, or for them to be fully recognized members of both at one and the same time. For some years, industry must be prepared to lose more than it may appear to gain'.

GROUP CHAIRMEN'S ADDRESSES

The Institution of Electrical Engineer
Abstract No. 254
Mar. 195

NORTH-WESTERN RADIO AND TELECOMMUNICATION GROUP: CHAIRMAN'S ADDRESS ©

By K. J. BUTLER, Associate Member.

'THE APPLICATION OF ELECTRONIC TECHNIQUES TO TEXTILES'

(ABSTRACT *of Address delivered at* MANCHESTER, *9th October*, 1957.)

The textile industry is still a basic industry of this country, and in Manchester—a city which has developed around its textile trade—16·4% of the total working population is engaged in textile pursuits.

Electronic techniques play an important part in the efficient production of modern textiles. In describing the way in which these techniques are applied, the subject has been divided as follows:

(*a*) Electronic control applications of a conventional nature, in which no particular problems arise due to the physical properties of textile materials.
(*b*) Electronic textile testing equipment.
(*c*) Control applications involving parameters peculiar to textiles, resulting in the use of specially developed transducers.

Conventional applications include level control of fibre in hopper, winding and reeling of yarns and fabrics, tension control during cloth-finishing processes, and edge alignment of fabrics during finishing.

The quality of textile goods, particularly yarns, is still largely controlled by constantly testing samples taken at various stages in the manufacturing process. There are, as yet, few examples of electronic control devices installed as part of the manufacturing process to maintain a desired quality of product. However, electronic techniques have made a considerable contribution to improved textile testing equipment, and the modern quality-control laboratory utilizes electronic instruments to measure the strength and thickness of fibres and yarns. The regularity of yarn thickness and strength is of paramount importance to the production of high-quality fabrics, and electronic devices have been developed to measure regularity and to analyse irregularities.

Some properties of fabrics, such as 'lustre', 'cover' and 'handle', are not so readily assessed quantitatively, but electronic techniques have sometimes enabled a more reliable quantitative assessment to be achieved.

Historic and economic considerations often prevent the application of electronic automatic control to textile processes. In some of these cases, an alternative is to use electronic equipment to monitor the correct adjustment of existing machinery, and instruments have been developed for this purpose.

Where large quantities of yarn or fabric are handled by one machine, we can justify automation. One example of this occurs in the finishing of cloth, where a specially developed electro-optical transducer is incorporated into a system controlling the correct alignment of the weft threads with respect to the warp threads.

Mr. Butler is with the British Rayon Research Association.

The Institution of Electrical Engineers
Abstract No. 2481
Feb. 1958

NORTH-WESTERN MEASUREMENT AND CONTROL GROUP: CHAIRMAN'S ADDRESS ©

By E. ROSCOE, J.P., Member.

'SPECIALIZATION IN EDUCATION AND INDUSTRY'

(ABSTRACT *of Address delivered at* MANCHESTER, *22nd October*, 1957.)

It appears to be appropriate for a Chairman of a Specialized Group to consider briefly the problem of specialization as it affects education, industry, and institutions such as ours. The problem of specialization has affected not only the electrical engineering industry, but many phases of society, e.g. the distributive trades in large centres of population, where each shop now specializes in a particular line. The hospital of the days of Florence Nightingale is rapidly disappearing. Special hospitals are built to deal with a specific disease. Doctors specialize more and more, and the general practitioner is rapidly becoming a diagnostician who treats only minor ailments.

Growth of Knowledge.—The designation 'electrical', 'mechanical' or 'civil engineer' is a fairly recent division, which only broadly describes the sphere of his professional knowledge and skill. The depth and breadth of knowledge expands at such a pace that it is difficult to keep abreast of it even in a very limited field. This can be seen by the number of technical publications in all languages. It is a problem to find time to read enough of them to feel fully informed in one's own field.

Within their lifetime men like Colonel Crompton and Dr. Ferranti saw the electricity industry develop from the very beginning into the highly specialized industry of to-day. It is remarkable that a biography of Madame Curie and the book on Calder Hall should have been published within the short space of 20 years. He would be a brave man who attempted to forecast the effects of recent scientific discovery on society during the next generation, especially now that man-made satellites are

Mr. Roscoe is with the North Western Electricity Board.

culating round us. If the growth of knowledge continues at present rate, more and more specialization appears to be evitable in industry, education and engineering institutions.

Education.—A balanced educational system has a twofold rpose: to teach the theory and practice of earning a living; and teach individuals about the business of living in a society.

Because of the scope of modern knowledge, specialization will, d must, occur at some stage of schooling. I sometimes nder whether professional institutions as a body take enough erest in education at the primary and secondary levels. It is this that a university education is built. At 11+ the child channelled, in the majority of cases, into one of three types of hool, which largely determines his future career. He must to a modern, a technical or a grammar school. At 15 years age the path diverges again to industry or to further educa-nal specialization. Again, at 18 or 19, when the student es to the university, he selects one or two fields of study. At e age of 22 or 23 he may enter industry and find himself once ore facing the problem of deciding in which branch to specialize. am convinced that the channelling of children at the stage of e 11+ examination is much too early in life. A period, from e age of 12 to 18 years, could be more usefully employed in a oader training than is given in any type of secondary school to-day and in guiding young people in their choice of a career. Special study should commence after the age of 18, continuing for at least three or four years at a university. A post-graduate course in social studies should be added to a degree or Higher National Certificate, if we are to aim at a well-balanced educational system. This might relieve some of the present tension in society. I believe that every engineer could benefit from a knowledge of the history of his profession and the lives of those who pioneered it.

Industry.—Since specialists must work together in teams in industry and at the university, and pool and co-ordinate their knowledge, residential courses covering a few months might be held in further educational centres, where problems in human relationships as well as technical problems can be dealt with. Part of the manager's duties is that of functioning specialist, especially where the work consists of scientific research and development, and to a lesser degree, the application of science to industry.

There is a danger that, if engineers do not give some thought to these and other related problems, the administrative side of the electrical engineering industry may pass into the hands of people who have not an engineering background and the engineer become merely a specialist in some branch of his profession.

DISCUSSION ON 'THE CONTROL OF NUCLEAR REACTORS'*

efore the MERSEY AND NORTH WALES CENTRE *at* CHESTER 10*th December*, 1956, *the* NORTH-WESTERN MEASUREMENT AND CONTROL GROUP *at* MANCHESTER 15*th January, and the* SOUTHERN CENTRE *at* PORTSMOUTH 3*rd April*, 1957.

Mr. J. A. Golder (*at Chester*): As a progression from their marks on control analysis by analogue-computation methods, ll the authors comment on the use of oscillators as a means of certaining the dynamic characteristics of a reactor? Can an alogy be drawn between reactor oscillators and harmonic alysers as employed for the determination of the control aracteristics of chemical plant?

Does the effect of the continual growth of fission products in a actor progressively decrease the reactivity of the system, or do me of the fission products contribute reactivity?

Mr. W. G. Proctor (*at Manchester*): In Section 5 the authors ention the 2-group theory of neutron emission for a nuclear actor, but advise caution in the interpretation of results based it. I recently investigated the servo problem involved in ntrolling a swimming-pool type of reactor under constant-wer conditions. At the early stages of such an investigation, act values of the transfer functions of the controlling elements nnot be known and some simplification is necessary. On the sumption of 2-group theory of neutron emission, eqns. (13) d (14) can be linearized by the substitutions $n = n_0 + \delta n$ and $= r_0 + \delta r$, where n_0 and r_0 are the neutron population and ncentration of precursors at a given reactor power level and and δr are small variations about this point. This removes e non-linearity on the right-hand side of eqn. (13) inherent in $(K - 1)$, since $(K - 1)$ may be a function of time. The uations can now be manipulated using the Laplace transform give a transfer function usable for constant-power conditions. a wide variety of control problems this reduces to a function the form 'proportional plus integral'. This is a very great plification on the original kinetic equations. Do the authors e that, for such investigations, a 2-group theory is adequate, would it be advisable to use a 3-group theory?

It is stated in Section 3.4 that a 25% safety margin is adequate for the invested reactivity of a control rod. The use of a safety margin would indicate that there is some uncertainty in the measurement of the leakage reactivity. How accurately can such a term be predicted?

Mr. M. W. Jervis (*at Manchester*): When running at nominally constant power, some reactors exhibit fluctuations in the neutron flux, commonly referred to as 'noise'. The amount of noise and its frequency distribution affect the design of a servo system intended to keep the flux constant. Will the authors give some details of the origin and nature of this noise in reactors?

The authors mention period trips used to shut down the reactor in the event of excessively short period. These are presumably positive periods, but reactors have been described which are fitted with negative period trips. What is the opinion of the authors on the desirability of such trips and in what circumstances might they be useful?

New boron- and cadmium-loaded scintillation-counter phosphors have recently become available, and it is claimed that they are particularly sensitive to neutrons. Are these likely to make the use of scintillation counters in reactor instrumentation any more attractive?

Mr. R. E. B. Dawson (*at Manchester*): References have been made from time to time, in papers on reactor control and instrumentation, to the use of compensated ionization chambers for flux-measuring circuits. In view of the advantages such a chamber would appear to have, for use in the high residual γ-activity resulting from large pressure vessels, I should be interested to learn the present position with regard to their development in this country.

When considering the safety philosophy applied to power reactors it is important to remember that reliability of supply is of prime importance. This means that the safety circuits should be arranged so as to avoid inadvertent shut-down of the reactor

* Cox, R. J., and WALKER, J.: *Proceedings I.E.E.*, Paper No. 2068 M, March, 5((see 103 B, p. 577).

due to faults on the protective equipment itself. One method of doing this employs a 'two out of three' system for the complete safety circuit. This differs from the method employed at Calder Hall, for instance, where 'two out of three' protection is employed on individual protective circuits only. As an example, it is possible to have a shut-down at Calder Hall during the start-up period by maloperation of the period trips, any one of which can cause shut-down. In a reactor designed primarily for power production, this could cause considerable delay in putting the station on load and is extremely undesirable. Will the authors comment on this?

Finally, it is mentioned in the paper that the shut-down cooling should be driven from 'a guaranteed supply of the utmost reliability'. What constitutes such a supply? Do the authors contemplate large-capacity batteries, or would generator sets equipped with flywheels of the no-break type be satisfactory?

Mr. W. J. A. Wills (*at Portsmouth*): Will the authors enlarge on the instrumentation necessary to measure the power level of a reactor? I understand that this is accomplished by measuring the neutron loss outside the reactor. What precautions, if any, are necessary in siting the instruments to obtain accurate measurements?

Mr. O. Nourse (*at Portsmouth*): What is the reason for the rather unexpected negative temperature coefficient of reactivity in nuclear reactors? Is this designed to be so or is it fortuitous, since a positive temperature coefficient would make reactors run away very easily?

Mr. T. J. Wilkinson (*at Portsmouth*): Can xenon poisoning lead to reactor instability, either by increasing as the power is reduced, thus tending to reduce the power further, or vice versa?

Mr. J. H. Bowen also contributed to the discussion at Manchester, and **Mr. P. Andrews** to that at Portsmouth.

Messrs. R. J. Cox and **J. Walker** (*in reply*): Reactor oscillators have long been used as a method of measuring the nuclear characteristics of samples of different materials. Their modern counterpart (which we prefer to call 'reactor transfer-function analysers') are used, as Mr. Golder states, to determine the dynamic characteristics of reactors. They are closely analogous to the harmonic analysers used in chemical-plant control studies. In one form of the equipment some fuel or an absorber is moved to produce a sinusoidal modulation of reactivity. The 'in-phase' and 'quadrature' components of the resultant neutron-flux modulation are measured. This can provide data on the temperature coefficients of reactivity and, at the higher frequencies, on the delayed neutron fractions. By measuring the modulation of temperature at appropriate points data can be obtained on heat-transfer characteristics. Exploitation of these techniques is only just starting, but we expect them to be used extensively in the development of nuclear power stations.

What happens to the reactivity of a reactor as time proceeds depends very much on the type of fuel used and on the reactor design. All atoms produced by the fission of fissile material are neutron absorbers, so we would expect the reactivity to decrease progressively. However, when natural uranium is used as a fuel, some of the uranium 238 is transmuted by neutron reactions to higher nuclides, some of which (e.g. plutonium 239) are themselves fissile. In most natural-uranium reactors the reactivity initially rises and then decays, the time scale being very dependent on the reactor design.

We are interested in Mr. Proctor's method of producing a 2-group simplification of the reactor-kinetic equations. For the purpose intended, where small-amplitude excursions are considered, it is probably a very satisfactory system. However, it is difficult to foresee its limitations without studying the metho in some detail.

Most reactor designs are far too complex to allow rigorou mathematical analysis. Some simplified models must be use and these calculations are supplemented by experiments o simplified physical models. The safety margin allowed in th design of control rods reflects mainly the degree of simplificatio used in the models.

Mr. Jervis raises the question of reactor noise. One contribu tion to noise is due to a statistical fluctuation in the neutro population of a reactor which arises mainly from a statistica variation in the number of neutrons produced per fission. I most high-power reactors the main contribution is due to th random variations in the number of neutrons detected in th ionization chamber. Both of these sources are purely randon in nature and therefore relatively straightforward to calculate but some reactors can exhibit fluctuations in power level arisin from purely mechanical causes, e.g. the rattling of fuel element or the bending of control elements through the flow of coolant These fluctuations are difficult to calculate and, of course, neve produce 'white' noise.

Negative period trips have been used on some researc reactors, but they are usually provided for protection agains failure of experimental equipment. For instance, flooding of a experimental 'thimble' with light water could be dangerous an its first indication would be a sudden drop in reactivity of th reactor.

We have used neutron-sensitive scintillation counters in low power fast reactors to provide a high sensitivity in a smal volume. In thermal reactors, where neutron cross-sections ar much larger, they are most unattractive compared with boron coated ion chambers, because of their inherent short irradiatio life, sensitivity to γ-radiation and to temperature and the effect of induced activities.

Mr. Dawson will be pleased to hear that we have a number o designs of γ-compensated ionization chambers which provide a compensation factor of about 30, but we have found a need fo them only in the swimming-pool type of reactor. In the Calde Hall type of power reactor the use of a graphite 'thermal column' plus lead shielding provides a much better solution.

It would, of course, be more logical to provide a 'two-out-of three' system for all the instrumentation at Calder Hall, but th start-up equipment is used for only a few hours at very infrequen intervals, so the chance of a spurious shut-down due to thi equipment is quite small.

The degree of reliability required from a guaranteed supply i directly related to the consequences of its failure, so each cas must be considered on its merits.

Using neutron-flux measurements as a measure of reacto power as described by Mr. Willis is the usual practice, but i ignores the heat due to fission-product decay, and the reading themselves change with changes in neutron flux distribution They cannot give accurate readings of power level, but d provide a method of controlling the neutron fission process.

In reply to Mr. Nourse we would point out that the negativ temperature coefficient is partly due to the fact that material expand when temperatures increase, so making the reactor les compact and therefore less reactive. However, there are a number of phenomena involved and a positive temperatur coefficient is possible, although such reactors are not usuall built because of their inherent instability.

Mr. Wilkinson is quite correct in describing xenon poisonin as a destabilizing influence, and instability is theoreticall possible; but the power level necessary for it to occur is muc higher than that of any reactor yet built in this country.

.374.32 : 518

The Institution of Electrical Engineers
Paper No. 2365 M
June 1957

THE DESIGN OF THE CONTROL UNIT OF AN ELECTRONIC DIGITAL COMPUTER

M. V. WILKES, M.A., Ph.D., F.R.S., Associate Member, W. RENWICK, M.A., B.Sc., Associate Member, and D. J. WHEELER, Ph.D.

e paper was first received 24th January, and in revised form 5th March, 1957. It was published in June, 1957, and was read before the Measurement and Control Section, *5th November, 1957.)*

SUMMARY

he function of the control unit of an electronic digital computer is provide the sequences of pulses, which, when applied to the store, hmetic unit and other units of the machine, cause the orders of the gramme to be executed. The paper discusses a number of related ys in which a systematic and flexible design for a control unit may achieved. In one group of systems the order code is determined the arrangement of diodes in a diode matrix, and in another by the ropriate threading of wires through a matrix of ferrite cores. The t part of the paper is concerned with logical design, and the second t with the practical design of a system using a ferrite matrix.

(1) INTRODUCTION

The paper is concerned with the design of the control or uencing unit of an electronic digital computer. The work cribed has been done in connection with the design of a allel binary machine (Edsac 2)[1,2] in which both orders and mbers are placed in the same store. Orders are of the single-dress type, and are executed one after the other, in the sequence which they stand in the store, except when the sequence is oken by a transfer of control, which may be either conditional unconditional. Most of the discussion is, however, equally olicable to the design of control units for machines of other es.

The execution of an order usually involves more than one step. r example, during an addition, the first step is to form the sum the number to be added (the incident number) and the number eady standing in the accumulator (the resident number), and second step is to transfer this sum to the accumulator, where eplaces the number originally there. In this case the steps to carried out are known in advance, and are the same for all ditions. In the case of multiplication, however, the steps to carried out depend on the digits in the multiplier; these digits examined one by one and, if the particular digit being mined is found to be a '1', the multiplicand is added to the mber in the accumulator (this requires two steps as before), number in the accumulator is shifted one place to the right, d the next digit of the multiplier is examined. If the digit is nd to be a '0', the number in the accumulator is shifted hout the addition of the multiplicand taking place.

In a parallel machine, each step required for the execution of order is accomplished by the application of a pulse to a set gates; it is the function of the control unit of the machine to duce the sequences of pulses necessary for this purpose. In ny machines the design of the control unit has been arrived by semi-empirical methods, and the resulting circuits, however ective they may be, are complex and non-systematic, and uld require considerable alteration if any appreciable change e to be made in the order code of the machine. The object the work described here was to develop a design for a control t which would be simple in overall structure, and would be applicable whatever the order code of the machine might be. The underlying principles were first discussed briefly at a Conference at Manchester University in 1951[3] and described in more detail in a paper published in 1953.[4] At that time it was thought that diode matrices would be used for generating the pulses, and that alteration of the order code would involve repositioning of the diodes in the matrices. As will be explained in more detail in the following Sections, later developments have led to the use of matrices of ferrite cores; these are threaded with wires according to the order code required, and any change in the order code requires partial rethreading.

There is a certain analogy between the way in which the machine performs the individual steps required for the execution of an order, and the way in which it performs the individual orders required for the execution of a programme. This suggested the term 'micro-programme' for the former. Similarly, the term 'micro-programming' was suggested for the process of drawing up the description of the steps from which the wiring diagrams for the diode or ferrite matrices are prepared. When these terms were first used in this connection, they were thought of as applying to machines in which the matrices were wired permanently so that micro-programming was nothing more than a method of designing the control unit of a machine. The terms have, however, been applied elsewhere to describe a system in which the order code of a machine is not permanently wired in, but can be set up by the programmer to have any form he chooses. The possibility of doing this was considered briefly in Reference 4, but it did not at that time appear that there would be a very strong demand among machine users for such a facility. Since the design of the control circuits of the machine is the province of the design engineer, the use of the term 'micro-programming' in this connection has proved misleading; for this reason, and because of the ambiguity just referred to, the term will not be used in the present paper.

(2) LOGICAL DESIGN

(2.1) Diode Matrices

Fig. 1 shows a simple form of control system using diode matrices. The register R controls a decoding tree which may be thought of as a multi-outlet switch. An input pulse is routed to the particular outlet line that corresponds to the number standing in the register. The outlet leads pass from the tree into two diode matrices, A and B. The outputs of matrix A are connected to the various transfer gates in the arithmetic unit and elsewhere in the machine. The outputs of matrix B are connected through delay networks to the register R, and cause the number standing therein to be changed in readiness for the passage of the next input pulse into the decoding tree. The configuration of diodes in matrix B thus determines the sequence in which pulses appear on the outlet leads of the decoding tree, and hence the sequence in which the various transfer gates throughout the machine are supplied with pulses. As so far described, the outlet leads of the

e authors are at the Mathematical Laboratory, University of Cambridge.

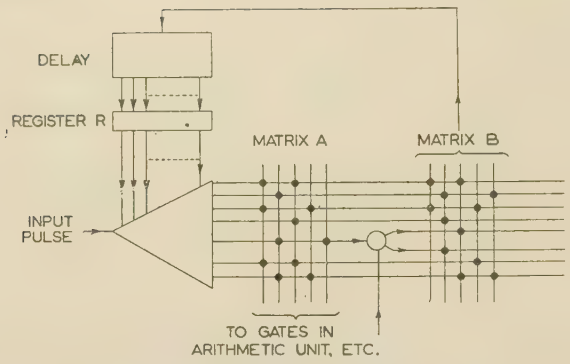


Fig. 1.—A control system using diode matrices.

decoding tree receive pulses in a fixed order. To make it possible to deal with operations in which certain of the steps are conditional, some of the outlet leads from the decoding tree are made to branch before passing into matrix B, one such branch being shown in Fig. 1. These branches are controlled by flip-flops in the arithmetic unit, or elsewhere, in such a way that the pulse from the decoding tree passes into one branch if the associated flip-flop contains a '1', and into the other branch if the associated flip-flop contains a '0'. Since the outputs from matrix B control the setting of the register R, it will be seen that the number which becomes set up on this register—and hence the next outlet lead of the decoding tree to receive a pulse—depends on the setting of the flip-flop. In the case of multiplication, the flip-flop concerned may be that containing the particular digit of the multiplier due to be sensed. In the analogy between the outlets of the decoding tree and orders in a programme, an outlet which branches before passing into matrix B corresponds to an order calling for a conditional transfer of control. If the outlet is made to branch before passing into matrix A, then the nature of the operation performed, as well as the sequence of operations, is conditional on the state of the flip-flop controlling the branch.

In addition to arithmetic registers, a digital computer must be provided with certain control registers, namely a register for holding the address of the order due to be executed after the execution of the current order, and a register in which a counting number can be held during the operations of multiplication, shifting, etc.; in addition, there may be one or more B-registers used for the automatic modification of instructions. The transfer gates controlling all these registers and their associated adder or adders can be operated by output pulses from matrix A in a manner similar to that in which transfer gates in the arithmetic unit are operated. In fact, the operation of the entire machine, including that of the store, can be brought under the control of a system such as that shown in Fig. 1. Both the extraction of an order from the store and the execution of that order can be included.

Once the sequence of operations required for the execution of a particular order has been initiated, the individual operations follow each other according to the configuration of diodes in matrix B. The correct sequence for a particular order is initiated by routing the function digits of the order (i.e. the digits which specify the operation called for by the order) to register R, where they occupy, say, the five or six least significant positions, the remaining positions having '0', transferred into them. When the register R has been set in this way, a pulse entering the decoding tree is routed to an outlet determined solely by the function digits of the order, and it is arranged that this outlet corresponds to the first of the sequence of operations called for by the order. It will be seen that no separate decoding tree is required for decoding the function digits of the order, nor is a register required for holding those digits during the execution the order.

An arrangement differing slightly from that of Fig. 1 is shov in Fig. 2. Here the outputs of matrix B are connected

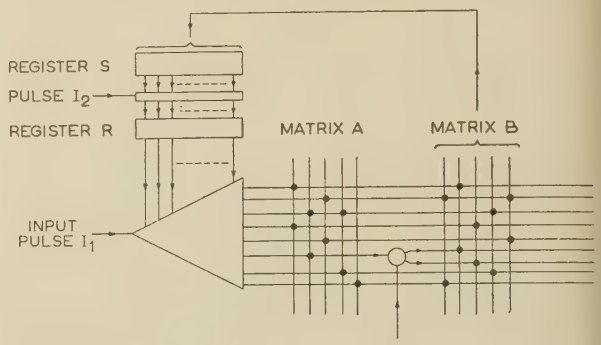


Fig. 2.—A modified version of the scheme shown in Fig. 1.

register S, and the number which becomes set up on this regist is later transferred to register R by the action of a pulse I_2 appli in parallel to a set of gates connecting the two registers. must occur after a suitable time interval from the application the input pulse I_1 to the decoding tree, and sufficiently in advan of the next pulse to be applied to the decoding tree.

Both systems so far described include provision for buff storage for the number which eventually becomes set up register R; in the system shown in Fig. 1 this buffer storage provided by means of a set of delay units, and in the system show in Fig. 2 by means of a separate register. The provision of buff storage is an essential logical requirement for the correct functio ing of any system in which a number set up in a register durir one step must be replaced by a new number before the next st can take place. In the system shown in Fig. 3, however, t necessity for buffer storage is avoided, and this system is, ther fore, theoretically capable of a higher speed of operation. T decoding tree is divided into two parts, each providing half t original number of outlets. There is a single A-matrix, joint controlled by the two decoding trees, but two separate B-matric are provided, one for each decoding tree; the outputs of o B-matrix are used to control the register of the decoding tr feeding the other B-matrix, and vice versa. During operatio pulses are applied alternately to the inputs of the two decodi trees. The advantage of faster operation with this system is some extent offset by the loss of flexibility resulting from the fa that the steps needed to execute an order must be divided equal between the two halves of the system.

(2.2) Ferrite Matrices

Diode matrices can be used economically in the control unit a machine with a relatively simple order code, but if the ord code becomes elaborate, including, for example, orders for floa ing binary or floating decimal operation, the number of diod required becomes inconveniently large. Accordingly, a syste using matrices of ferrite cores with coincident-current selectio has been developed.

A diagram showing the principle of operation of the syste is given in Fig. 4. For simplicity, a 3×3 matrix only shown. The matrix is controlled by two registers, X and consisting of flip-flops; these flip-flops are connected throu gates to driving valves, which are in turn connected to the ro and column circuits of the matrix. In the diagram, the row a column circuits are shown as though they consisted of straig wires passing through the cores, although in practice they a more likely to consist of multi-turn windings connected in serie

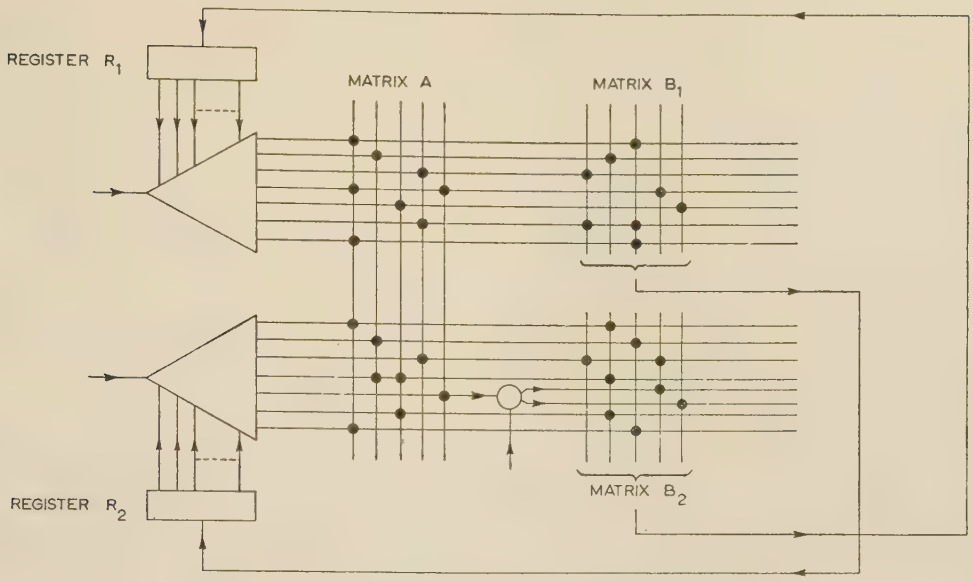


Fig. 3.—An alternative control system using diode matrices.

addition to the coils forming part of the row and column rcuits, the cores carry bias windings which are connected in ries, and through which a steady current is passed. The bias ndings are not shown in the diagram. In the experimental atrices 8 mm ferrite cores have been used.

Normally only one flip-flop in the X-register and one in the -register are set, the others being in the reset condition; no rrent passes through the row and column circuits, and all cores e magnetized to saturation in the (say) negative direction. hen a drive pulse is applied to the input shown in the diagram, e gates connecting the flip-flops with the driving valves are ened, and current passes through one of the row circuits and e of the column circuits. These currents are of such magnitude at their combined effect is sufficient to reverse the direction of agnetization of the core standing at the intersection of the lected row and column, but insufficient to reverse the magnetiza- n of other cores on the same row and column. At the termina- n of the drive pulse, the selected core returns to its original ate of magnetization.

Two independent sets of wires are threaded through the cores cording to the requirements of the order code of the machine. hen a core of the matrix is switched by the application of the ive pulse, an electromotive force is induced in all the wires ssing through that core, and when the core returns to its rmal state an e.m.f. of opposite polarity is induced. One end each wire is earthed, and the other is connected to a non-linear plifying circuit, so designed that an output corresponding to e second of these two induced e.m.f.'s only is obtained. One the two sets of wires threaded through the cores corresponds to atrix A of Fig. 2, and the outputs from these wires are used to erate transfer gates in the arithmetic unit and in various other its of the machine.

The second set of wires corresponds to matrix B, and the tputs from these wires are connected to the set or reset inputs e flip-flops in the X- and Y-registers. The wires are threaded rough the cores in such a way that, when any particular core turns to normal after having been selected, it causes the two p-flops, one in register X and one in register Y, which were ig nally in the set condition to be reset, and two other flip-flops be set. The application of a succession of drive pulses thus es a succession of cores in the matrix to be switched in a scribed order. If the selection of one core is to be followed

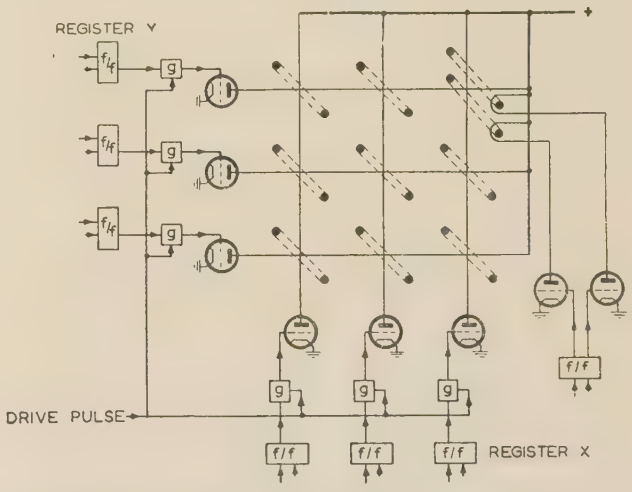


Fig. 4.—A control system using a ferrite matrix.
The bias winding and the output windings threading the cores are omitted.

by the selection of another core in the same row or column, the corresponding flip-flop in the Y- or X-register may remain set throughout. It will be observed that, since use is made of the output obtained from a core when it returns to its original condition (rather than the output obtained when it is selected), it is not necessary to provide buffer storage for the numbers to be set up on the X- and Y-registers. Storage is logically necessary, but it is, in fact, provided by the selected core itself.

It will be seen that the registers X and Y perform a function similar to that performed by register R in Fig. 1. It has been pointed out, however, that, at any given time, no more than one flip-flop of each of the registers X and Y is in a set condition, the others being in the reset condition; thus, $2m$ flip-flops are required to control m^2 cores in a square matrix, compared with the n flip-flops required to control the 2^n outputs of the decoding tree of Fig. 1.

For reasons explained earlier, it is necessary that some of the steps in the operation of the control unit should be dependent on information supplied from outside that unit. This is made possible by locating two cores, instead of one, at some of the

intersections of the matrix. Each of these cores has an extra winding in addition to the three previously mentioned. At any given time, a current flows through the extra winding on one of the cores, and this is sufficient, in conjunction with the current flowing through the normal bias winding, to prevent the core from being switched even when the intersection containing it is selected. The matrix behaves, therefore, as if only the other core were present at the intersection. Thus the action which takes place when an intersection with two cores is selected is dependent on which core is biased off. In Fig. 4 one intersection with two cores is shown, and the action which takes place is dependent on the setting of the flip-flop controlling the currents passing through the extra windings.

It is possible to locate more than two cores at an intersection and obtain a multi-way conditional action; for example, four cores, each with two windings in addition to the three ordinarily required, can be used to provide a four-way conditional action, controlled by the settings of two flip-flops. Any flip-flop, or pair of flip-flops, in the machine may be connected to control any conditional intersection. Where it is possible, however, economy in driving valves may be secured by connecting the conditional windings on cores at a number of intersections in series, and using the same flip-flop to control them. The controlling flip-flops may either be flip-flops existing in the registers of the machine, or special flip-flops, introduced for the purpose and set, as appropriate, through gates operated by outputs of the control matrix.

An alternative way of obtaining conditional action is illustrated in Fig. 5. One or more of the rows of the matrix is made double,

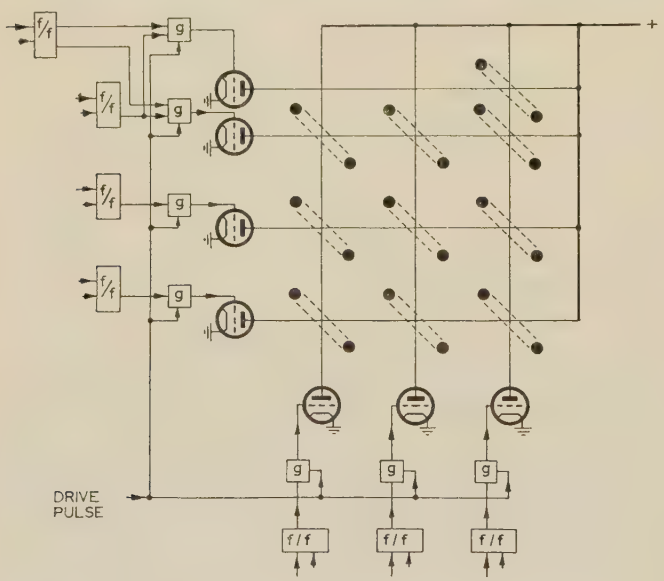


Fig. 5.—An alternative system for obtaining conditional action.

and the branch which receives a pulse when the drive pulse is applied (assuming that the relevant flip-flop of register Y is set) is determined by the setting of a flip-flop. If more than two branches are provided, multi-way conditional action may be obtained. Conditional action at more than one intersection on the same row, and controlled by the same flip-flop, may be conveniently provided in this system, but if it is required to control independently a large number of intersections the provision of complicated switching circuits becomes necessary.

The basic scheme shown in Fig. 4 may be modified by using magnetic cores instead of flip-flops in the X- and Y-registers. A diagram of one stage of a register using magnetic cores is

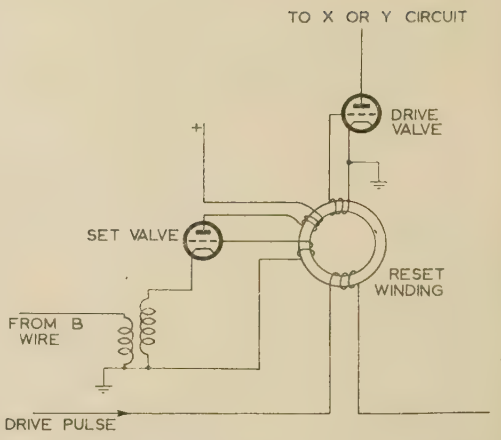


Fig. 6.—One stage of an X- or Y-register composed of magnetic cores.

shown in Fig. 6. At any given time all the cores, except one in each register, are magnetized in a certain direction and will be said to be in the *reset* condition. The remaining cores are magnetized in the opposite direction, and will be said to be *set*. The reset windings of all the cores in both registers are connected in series, and the application of the drive pulse causes a current to pass through the circuit so formed. This causes the cores which are in the set condition to be reset, and pulses to be supplied to the valves driving one of the rows and one of the columns of the matrix, with the result that a selected core is switched. This core is threaded by a pair of B-wires and, when it returns to normal, pulses induced in these B-wires cause one of the cores in the X-register and one in the Y-register to be set ready for the next operation. It will be noticed that positive feedback is applied to the setting valve associated with each core.

Either method of providing conditional action can be used with the scheme of Fig. 6, but the second has the advantage that the flip-flop, and the associated switching circuits controlling the conditional action, can be replaced by magnetic cores. The scheme shown in Fig. 6 is due to Mr. C. H. Lindsay, who has demonstrated a control system based on it.

In the systems based on diode matrices described earlier, it was possible to avoid the use of special equipment for decoding the function digits of the order, by placing the function digits themselves in the register R at the beginning of the operation. This is not possible when ferrite matrices are used, and the provision of some special decoding equipment is essential. Perhaps the most straightforward method would be to provide circuits which would cause the appropriate flip-flops (or cores) in the X- and Y-registers to be set according to the proper starting point for each order. Later in the paper a system is described in which part of the decoding is performed in a special section of the control matrix itself.

There are a variety of alternative coincident-current selection systems which may be used for the control matrix. Some of these depend on coincidence between currents in three or more windings, and lead to economy in the number of flip-flops and driving valves, but have the disadvantage of requiring more windings on each core. They have the further disadvantage that, with a given number of cores in the matrix, each driving valve must supply current to more windings in series. Since the number of windings which can be thus driven is the factor which ultimately determines the maximum size of a matrix, it follows that such systems do not lend themselves so well to the use of large matrices as the system described earlier. An interesting scheme depending on twofold coincidence only is shown in Fig. 7. Here a triangular matrix of 10 cores is controlled by a

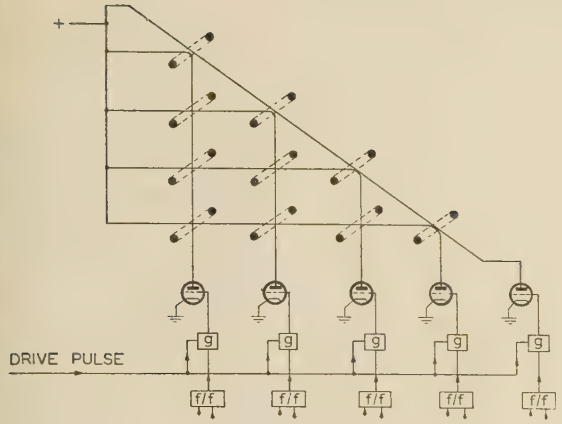


Fig. 7.—A triangular control matrix.

gister of five flip-flops, two of which are set at any given time. ne scheme can obviously be extended to larger matrices and, as mpared with that using a square matrix, results in an economy f flip-flops and driving valves. It does, however, suffer from e disadvantage just mentioned of requiring a relatively large mber of windings to be driven in series, and would not therefore as suited to the construction of a large system as a scheme sed on a square matrix.

(3) LAYOUT OF LARGE MATRICES

Experience has shown that matrices containing up to at least rows and columns are quite practicable (see Section 5.2). If e scheme of Fig. 4 is used, however, some attention must be ven to the layout of the cores. A geometrical arrangement of e cores on the lines of that shown in the Figure, in which one, o or four cores are placed at each intersection, cannot be mployed very conveniently for the construction of an actual atrix. An alternative system will therefore be described in hich the cores lie physically at the intersections of a rectangular id, but are connected electrically in a manner which corresponds that shown in Fig. 4.

The matrix may be regarded as being formed of a number of b-matrices, each containing 16 cores; there are two types of b-matrix, A and B, and they are arranged in a chessboard ttern. Fig. 8 shows the wiring of the two types of sub-atrix (omitting bias and conditional windings), and Table 1

Table 1

A	B	A	B	A	B	A	B
B	A	B	A	B	A	B	A
A	B	A	B	A	B	A	B
B	A	B	A	B	A	B	A
A	B	A	B	A	B	A	B
B	A	B	A	B	A	B	A
A	B	A	B	A	B	A	B
B	A	B	A	B	A	B	A

ows how 32 of each type may be used to form a matrix with 2 rows and 32 columns. It is to be understood that where the dge of one sub-matrix abuts on the edge of an adjacent sub-atrix, the four wires emerging from one sub-matrix are con-ected to the four wires entering the other. If the wires passing om left to right and from top to bottom through the arrange-ent of sub-matrices shown in Table 1 are traced it will be found at the entire arrangement is equivalent to a matrix of straight ires having four cores at 64 of its intersections, two cores at 2 of its intersections, and one core at 512 of its intersections; e are no cores at the remaining 320 intersections. The fact

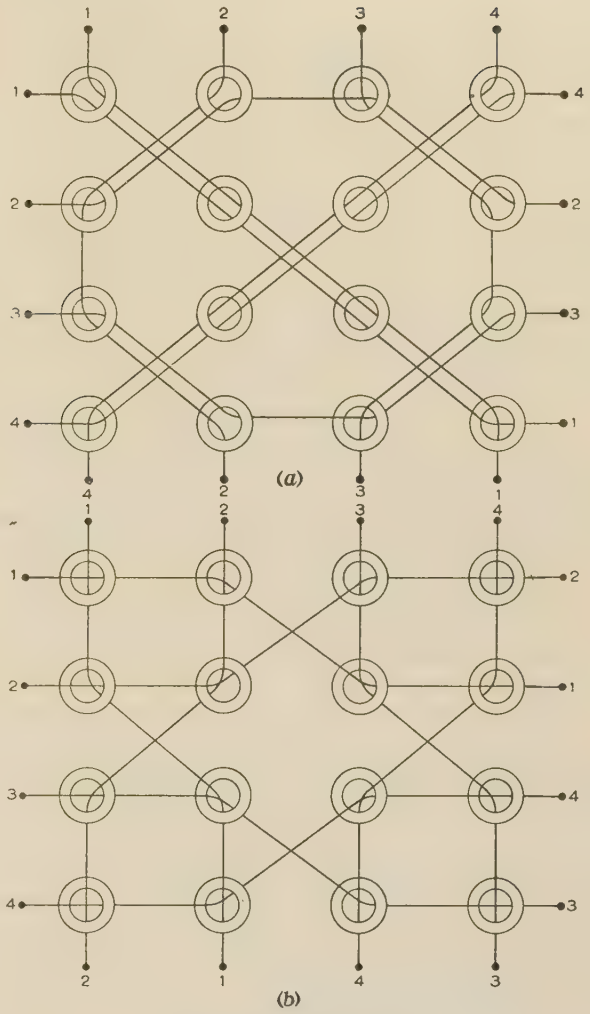


Fig. 8.—Sub-matrices.
(*a*) Type A. (*b*) Type B.

that there are no cores at some intersections means that certain otherwise valid settings of the X- and Y-registers do not correspond to steps in the control sequence; on the other hand, the number of windings connected in series is the same for each row and for each column, and, moreover, the three types of intersection are distributed uniformly with respect to the rows and columns. If a number of intersections smaller than that provided by the arrangement of Table 1 is sufficient for a particular application, one or more of the rows or columns of sub-matrices may be omitted. The resulting matrix will still have the same number of cores on all rows and on all columns, although the distribution of the various types of intersection may no longer be uniform.

If the scheme of Fig. 5 is used, a proportion of single, two-way and four-way intersections similar to that provided by the arrangement just described may be obtained in the manner shown in Fig. 9. The arrangement there shown may be regarded as consisting of a matrix with 11 rows of 16 columns; the top row is split into four branches, the next two rows each into two branches, and the last 8 rows remain single. There are thus 16 four-way intersections and 128 single intersections. These are the same numbers as would be provided by a matrix of 16 rows and columns composed of A and B sub-matrices in the way described above. The scheme of Fig. 9 suffers, however,

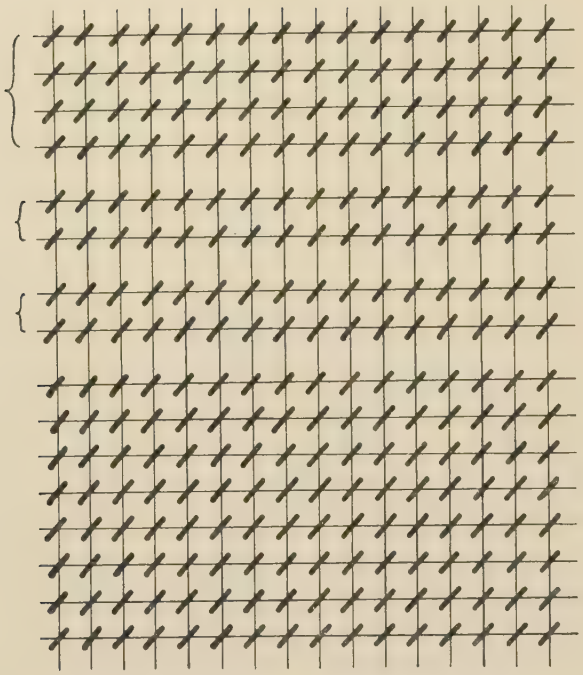


Fig. 9.—Possible arrangement of cores using the system shown in Fig. 5.

from the disadvantage that all the multiple intersections must be brought under the control of not more than four flip-flops.

(4) FUNCTION DECODING

It has already been mentioned that special arrangements are necessary to ensure that the correct intersection is selected at the beginning of each order. It would be possible, though not very convenient, to make use for this purpose of a sequence of four-way conditional intersections, controlled by flip-flops on which the function digits are set up. Another method would be to rely on the use of decoding circuits based on the use of diodes for setting the correct flip-flops in the X- and Y-registers. Many schemes based partly on the use of diodes and partly on the use of cores are possible. The one described below was devised for a machine in which seven function digits are contained in each order.

The seven function digits are decoded in two stages. The first stage is performed in an ordinary four-way conditional intersection in the control matrix; this intersection, which is controlled by flip-flops on which the first two function digits of the order are set up, is selected at the beginning of the execution of any order. The second stage takes place in a special section of the control matrix containing four columns each with 32 cores. These cores each have two conditional windings, but no row windings, double the usual current being passed through the column winding when a particular column is selected. The conditional windings are connected to driving valves controlled by the outlets of two conventional diode decoding trees, one of which is fed with the third, fourth and fifth function digits, and the other with the sixth and seventh function digits. The connections are shown diagrammatically in Fig. 10.

When one of the four cores at the four-way intersection in the main part of the control matrix is selected at the beginning of the execution of an order, a pulse induced in a B-wire threading it causes the flip-flop controlling one of the columns of the decoding section of the matrix to be set. When the next drive pulse is

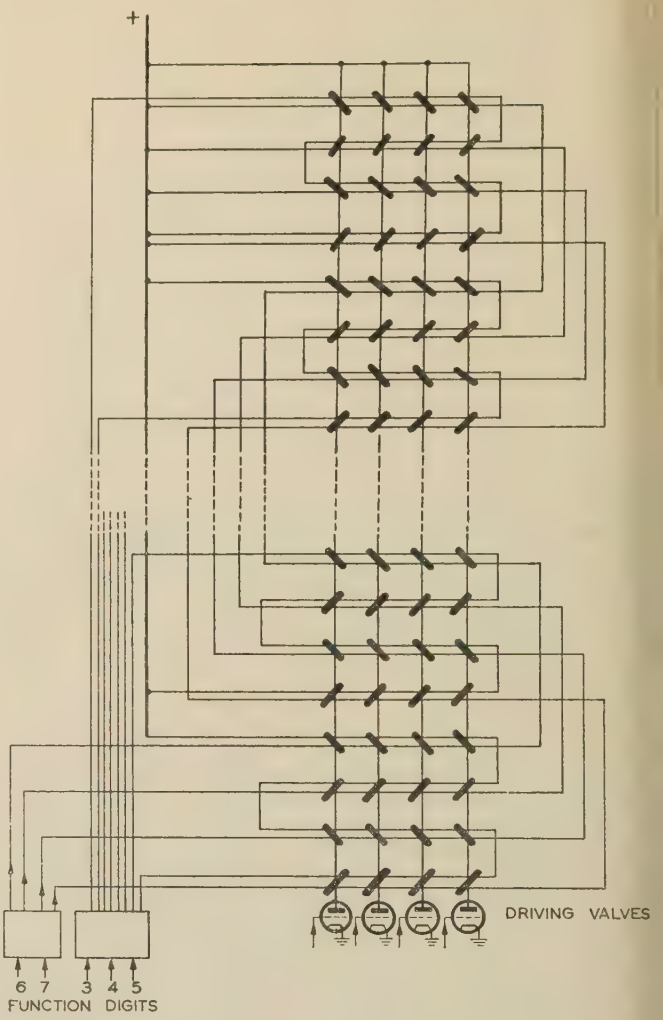


Fig. 10.—Function decoder.

applied, one of the 32 cores on that column is switched, according to the values of the five function digits feeding the decoding trees. The switching of this core causes the first step in the execution of the order to be performed, and initiates the performance of subsequent steps in the usual manner.

(5) ENGINEERING DESIGN

The satisfactory operation of the scheme of Fig. 4 requires (*a*) that the voltage induced in an output wire when a core is returned to its normal condition by the bias current, after having been selected, should be greater than a certain minimum value, and (*b*) that the break-through voltage, induced in an output winding when a core which has been disturbed by one drive current only is returned to normal, should be less than a certain maximum value. The magnitude of the break-through voltage determines the maximum number of cores on the same row or column which can be threaded by a given output winding. The operating conditions should be chosen so as to make the tolerances on the drive current, I_D, and the bias current, I_B, as large as possible.

(5.1) The Practical Design of Ferrite Matrices

It has been found experimentally that when a rectangular-loop ferrite core is fully switched by a current pulse, the maximum

ate of change of flux (or peak output voltage) increases with the pplied magnetomotive force. As a corollary to this, since the lux change in the core is constant, the switching time decreases vith the applied m.m.f.

The m.m.f. in a selected core operating in the matrix of Fig. 4 s proportional to the sum of the row and column drive currents ess the bias current, and is thus obviously a minimum when oth drive currents are at the low extreme of their tolerance ange and the bias current is high. Under these conditions, if and q are the fractional tolerances permissible in drive and bias urrents respectively, we must have

$$2I_D(1-p) - I_B(1+q) > I_S$$

where I_S is the current required to switch the core in a time orresponding to the length of the drive pulse. If I_B is nominally qual to mI_D the condition becomes

$$2I_D(1-p) - mI_D(1+q) > I_S$$

whence $$m < 2(1-p)/(1+q) - I_S/I_D(1+q) \quad . \quad . \quad . \quad (1)$$

Since the output voltage (obtained when the core returns to its normal condition) is proportional to the bias current, we must have

$$mI_D(1-q) > I_0$$

where I_0 is the current required to give the minimum permissible output voltage,

or $$m > I_0/I_D(1-q) \quad . \quad . \quad . \quad . \quad . \quad . \quad (2)$$

For conditions (1) and (2) to be satisfied simultaneously we must have

$$I_D > [1/2(1-p)][I_S + I_0(1+q)/(1-q)] \quad . \quad . \quad (3)$$

If, instead of being allowed to vary independently, the bias current is made proportional to the sum of the drive currents, condition (1) is replaced by

$$m < 2 - I_S/I_D(1-p) \quad . \quad . \quad . \quad . \quad . \quad . \quad (4)$$

and condition (2) by

$$m > I_0/I_D(1-p) \quad . \quad . \quad . \quad . \quad . \quad . \quad (5)$$

From conditions (4) and (5) we obtain

$$I_D > (I_S + I_0)[1/2(1-p)] \quad . \quad . \quad . \quad . \quad (6)$$

Since the application of a drive current in one row or column only must not begin to switch the core, a further condition must be satisfied. If I_K is the value of current at which the core begins to switch, we must have, in the case when the bias and drive currents vary independently,

$$I_D(1+p) - mI_D(1-q) - I_K < 0 \quad . \quad . \quad . \quad . \quad (7)$$

When the bias current is proportional to the drive current, the condition becomes

$$I_D(1-m)(1+p) - I_K < 0 \quad . \quad . \quad . \quad . \quad (8)$$

As long as conditions (7) and (8) are satisfied, the break-through voltage increases with increasing I_D, and, since the residual permeability of the core increases as the operating point moves toward the knee of the B/H loop, the more negative the left-hand side of (7) or (8) is, the greater are the operating margins and the smaller the break-through voltage. Thus, to minimize the break-through voltage the minimum value of I_D and the maximum value of m which satisfy all the above conditions should be chosen. An added advantage which is obtained by reducing I_D is that the power input required for the matrix is decreased. If these considerations are borne in mind, it will become apparent, when conditions (3) and (6), and (7) and (8) are compared, that making I_B proportional to the sum of the drive currents is advantageous. As an example we may take $p = q = 0{\cdot}1$, $I_S = 3{\cdot}6$ AT and $I_0 = 6{\cdot}4$ AT. Then, in the case where I_B and I_D vary independently, $I_D > 6{\cdot}35$ AT. Taking $I_D = 6{\cdot}5$ AT we find $1{\cdot}14 > m > 1{\cdot}09$. With $m = 1{\cdot}1$ we find that I_K must be greater than $0{\cdot}72$ AT if the core is not to be switched by one drive current under the most unfavourable circumstances. On the other hand, when I_B is proportional to I_D we obtain $I_D > 5{\cdot}5$ AT. With $I_D = 6$ AT we find $1{\cdot}33 > m > 1{\cdot}18$ and thus we may take $m = 1{\cdot}25$. With these values for I_D and m, I_K must be greater than $-1{\cdot}35$ AT, and, since I_K is always positive, it follows that the operating margins are considerably increased in this case. Even when $q = 0$, i.e. when no tolerance is permitted on the bias current, we find that, with $I_D = 6$ AT and $m = 1{\cdot}25$, I_K must be greater than $-0{\cdot}9$ AT, and therefore the margin is still less than in the case where I_B is proportional to I_D.

The behaviour of a typical core, as I_D and I_B are varied, can be seen in Fig. 11, where contours of constant output voltage and of constant break-through voltage are plotted against I_D and I_B. When I_D and I_B vary independently, the operating region in Fig. 11 is a rectangle. The rectangle shown corresponds to the case in the example above. If I_D and I_B vary together, the operating region becomes a straight line passing through the point $I_B = I_D = 0$, with a slope equal to $1/m$. The straight line through the origin shown in Fig. 11 corresponds to $m = 1{\cdot}25$, and shows the improvement obtained in the ratio of minimum V_0 to maximum V_b for the same tolerance variation as in the case of independent variation of I_B and I_D.

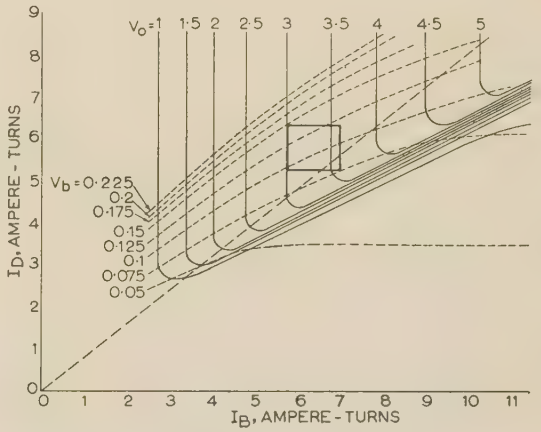


Fig. 11.—Experimental data for 8 mm ferrite cores.

I_D and I_B can be made to vary together by using the system shown in Fig. 12. The drive valves are arranged with a common cathode resistor to a negative return voltage, V_1 and V_{1a} being provided to accept the current when none of the drive valves are conducting. The current flowing in the bias winding is the sum of the row and column drive currents, and the correct ratio of I_B to I_D is obtained by adjusting the relative number of turns on the bias and drive windings. In the matrix developed for Edsac 2, for example, the drive winding has 40 turns and the bias winding 25 turns, giving $m = 1{\cdot}25$.

The choice of dimensions and number of turns for the cores composing the matrix, and of the type of valve to be used for driving, involves a compromise between a number of conflicting requirements. Since it is desirable that a given output winding should thread each core a few times only, a large output voltage per turn, and hence a core with a large cross-sectional area, is

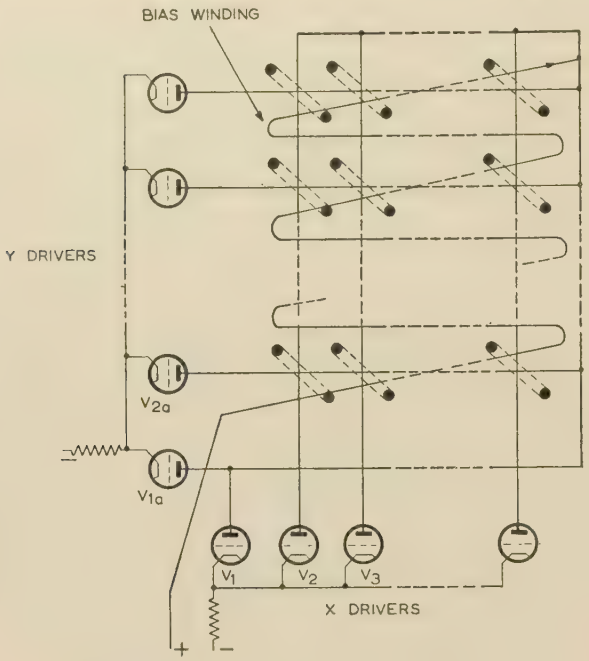


Fig. 12.—Connection of bias winding.

indicated; the available winding space also increases with the size of the core. However, if the cores are too large it may not be possible for the selected core to be switched sufficiently fast, or the size of the driving valves required may become prohibitive. Moreover, the temperature rise in a core subjected to continual switching increases with the dimensions of the core. This temperature rise may set the limit to the number of times per second that it is possible to switch a given core.

The number of turns on the drive windings should be made as small as is possible, consistent with satisfactory switching by the drive valves used, in order to reduce leakage inductance and to make the time of switching as short as possible. The maximum number of cores on each row or column which can be driven by one valve is determined by the voltage developed across the selected core, the total leakage inductance of the unselected cores, and the h.t. supply; the latter is in turn limited by the permissible anode dissipation of the driving valves. A secondary limitation on the number of cores in a row or column is imposed by the delay and distortion in the leading edge of the current pulse, which are due to leakage inductances and stray capacitances. With a drive current of 150 mA an 8 mm core can be operated on the straight line of Fig. 11 if the drive windings are composed of 40 turns each, and the bias windings of 25 turns. Experiments have shown that, if a column of 32 cores is operated under these conditions, the selected core is switched in less than one microsecond. The drive current of 150 mA can be provided by a 12E1 valve with a 150-volt h.t. supply.

(5.2) Break-through

The maximum number of cores through which it is possible, in a practical system, to thread a given output winding depends on the ratio of the output voltage obtained when a core is selected to the break-through voltage obtained when it is half-selected. Discrimination between wanted and unwanted signals can be made easier by increasing the number of times an output wire threads each core through which it passes, thus increasing the absolute level of both signals. This, however, increases the leakage inductance in series with the output winding, and in a large matrix, in which one output winding may thread up to 80 cores, this increase may be serious if more than two or three turns are used. A limitation is also imposed by the winding space available, since there may be 12 or more output windings threading a single core.

It is possible to reduce the break-through signal by cancellation, using extra cores for the purpose. In the simplest system, one core having a single drive winding only is included in each row and in each column; such a core will be half-selected each time the appropriate row or column is selected. Each output winding is threaded in a negative direction through the extra core the same number of times as it is threaded in a positive direction through other cores on the same row or column. This method of cancelling break-through is quite effective, but again has the disadvantage of increasing the leakage inductance in series with the output winding.

Measurements made on every core of a small test matrix of 8×6 cores showed that the break-through voltage from an unselected core was less than 0·1 volt per turn, and that the output voltage from a selected core was greater than 3 volts per turn. The operating speed was such that the time interval between the switching of successive cores was 1·5 microsec. Similar measurements made on a matrix of 3×3 cores, with the central row and the central column extended symmetrically to contain a total of 32 cores, indicated that similar figures were to be expected in the case of a complete matrix of 32×32 cores operating at the same speed. In the light of these figures it was decided that cancellation would be unnecessary in the case of the matrix for Edsac 2, provided that an upper limit could be placed on the number of times a given output winding threaded cores on the same row or column. By choosing carefully the sequence in which the cores would be switched during the operation of the matrix, no serious difficulty was encountered in making this limit equal to four, so that the total number of cores which could contribute to the break-through voltage was eight. Due allowance was made for the fact that a conditional core subjected to a drive current of $2I_D$ contributes about twice the normal break-through voltage contributed by a half-selected core. The output windings thread each core three times, giving a minimum output signal of 9 volts.

(6) CONCLUSION

The work described in the paper was undertaken primarily (as has already been stated in Section 1) with a view to designing a control system for Edsac 2. This computer, which is intended ultimately to have a comprehensive order code and a correspondingly large control matrix, is at present operating with a restricted order code provided by the 8×6 matrix referred to in Section 5. The control unit has been operating with this matrix for a year, and its performance has been very satisfactory.

(7) REFERENCES

(1) Renwick, W.: 'Edsac II', *Proceedings I.E.E.*, 1956, **103 B**, Suppl. 2, p. 277.

(2) Renwick, W.: 'A Magnetic Core Matrix Store with Direct Selection using a Magnetic-Core Switch Matrix', *Proceedings I.E.E.*, Paper No. 2236 R, October, 1956 (**104 B**, Suppl. 7).

(3) Wilkes, M. V.: 'The Best Way to Design an Automatic Calculating Machine', Report of Manchester University Computer Inaugural Conference, July, 1951, p. 16 (Manchester, 1953).

(4) Wilkes, M. V., and Stringer, J. B.: 'Micro-Programming and the Design of the Control Circuits in an Electronic Digital Computer', *Proceedings of the Cambridge Philosophical Society*, 1953, **49**, Part 2, p. 230.

[The discussion on the above paper will be found on page 144.]

.1.374.32

The Institution of Electrical Engineers
Paper No. 2389 M
July 1957

A DECIMAL ADDER USING A STORED ADDITION TABLE

By M. A. MACLEAN, M.Sc., and D. ASPINALL, B.Sc.

(*The paper was first received* 19*th March, and in revised form* 29*th April*, 1957. *It was published in July*, 1957, *and was read before the* Measurement and Control Section *5th November*, 1957.)

SUMMARY

A serial decimal adder is described which accepts numbers in binary-coded form. The binary digits, which are handled in parallel, are decoded into a set of pulses which actuate a built-in addition table covering all the possible sums. A special form of number representation has enabled the adder to be constructed more economically than would otherwise have been the case, while the advantages of this type of adder have been retained.

The circuits use square-loop magnetic cores for all logical functions and junction transistors as pulse amplifiers.

LIST OF SYMBOLS

F = Driving m.m.f. applied to core, ampere-turns.
N_0 = Number of turns on core output winding.
S = Constant relating drive m.m.f. to core switching time, ampere-turn-sec.
k = Constant related to the coercive force of the magnetic material, ampere-turns.
$\Delta\Phi$ = Total flux change in core between remanent states, webers.
Δt = Core switching time, sec.
V_b = Transistor base voltage when bottomed, volts.
I_b = Transistor base current when bottomed, amp.
V_c = Transistor collector voltage, volts.
I_c = Transistor collector current, amp.
c = Slope of the transistor I_b/V_b characteristic, mhos.
d = Intercept of the I_b/V_b characteristic on the voltage axis, volts.
V_{bias} = Base bias voltage, volts.

(1) INTRODUCTION

Despite the widespread use of binary numbers in digital computers, there are times when it is more convenient to use a decimal representation. A particular example is a computer designed for data-processing applications where the ratio of input-output time to computing time is large. In these cases the disadvantage of having to provide more complicated computing and storage circuits may well be offset by the fact that no binary-to-decimal conversion is required for input and output. For the same reason, the longer multiplication time for numbers expressed in decimal notation is not so important.

Although the representation of numbers may be basically decimal it is customary, for storage and computing purposes, to express each decimal digit as a binary number. The advantage of being able to use two-state circuits can then be obtained. Many types of binary coding can be used, the choice depending on the computing circuits which are to be employed.

(2) ALTERNATIVE FORMS OF ADDING CIRCUIT

Decimal adders have been described in which each decimal digit is represented by four binary digits in a straightforward manner, the addition being carried out serially from the point of view of the decimal inputs and in parallel from the point of view of the binary digits. Such a device is basically a 4-digit parallel binary adder.

The core-transistor circuit element is not very well suited to this type of addition as the signal delay due to the switching time of the transistor makes it difficult to arrange for the propagation of the binary carry. This problem can be surmounted if the addition is performed in two stages or if the four binary digits representing each decimal digit are handled serially instead of in parallel. In either case the speed of the adder is reduced.

Another approach is to use a matrix of coincidence gates, suitably connected, which stores all the possible sums. The input numbers are then applied to this matrix coded in a one-out-of-ten form. This method avoids the problem of the binary carries and is well suited to the core-transistor circuit element as all the operations can take place in strict synchronism with the basic repetition frequency of the system. The disadvantage is the large amount of equipment needed and the fact that much of it is lying idle at any time.

If the decimal digits are expressed in the scale of five, there will be two addition tables containing about a quarter of the number of gates required by the base-10 system. However, a base-5 carry will now be required. Although it would be easier to allow for this single carry rather than for the three which arise in parallel binary addition, the system does not realize all the potential advantages of the addition-table type of adder.

Another way of representing decimal digits will now be described which allows this type of adder to be constructed as economically as one using bi-quinary coding, but which has no internal carry.

(3) THE RESIDUE NUMBER SYSTEM*

If two positive integers a and b are divided by a third positive integer m, one obtains two remainders or residues r_a and r_b which lie between 0 and $m-1$. In the terminology of congruences,[1]

$$a \equiv r_a \text{ (modulo } m)$$

and

$$b \equiv r_b \text{ (modulo } m)$$

If the remainder on dividing $a+b$ by m is r_{a+b}, it is possible to show that

$$a + b \equiv r_a + r_b \equiv r_{a+b} \text{ (modulo } m)$$

and also that

$$ab \equiv r_a r_b \equiv r_{ab}$$

where r_{ab} is the remainder on dividing ab by m.

If the two integers a and b are now divided by a whole set of moduli $m_1, m_2, \ldots m_s, \ldots m_n$ in turn, a set of residues is obtained for each number which is characteristic of that number. The theorems quoted above then show that, if corresponding members of the two sets of residues are added together and the result is divided by the corresponding modulus, the set of residues so obtained is characteristic of the sum of a and b. An analogous

Mr. Maclean and Mr. Aspinall are in the Electrical Engineering Department, University of Manchester.

* This method of number representation was brought to the notice of the authors by Dr. A. Svoboda of the Academy of Sciences, Prague.

result holds for the product of a and b when corresponding members of their sets of residues are multiplied together.

It is not possible to represent an infinite number of integers uniquely by means of a finite number of moduli, since, as is shown in the Appendix, any set of residues represents a whole range of numbers differing by multiples of the lowest common multiple of the moduli. Provided, however, that attention is confined to a group of numbers lying between multiples of the l.c.m., a unique representation of this group can be obtained.

The important feature of this method of number representation in the present instance is that addition or multiplication is carried out by adding or multiplying corresponding members of two sets of residues without any operation being affected by the result of another, i.e. there are no carries. In particular, it provides a method of partitioning a decimal digit to obtain economy in a stored-table type of adder without thereby introducing the problem of a carry.

(4) REPRESENTATION OF DECIMAL DIGITS BY THE RESIDUE SYSTEM

The moduli used for representing a decimal digit are 5 and 2. Since their l.c.m. is 10, any numbers differing by a multiple of 10 will be represented by the same pair of residues. Thus it follows that when two decimal digits are added together, the least significant digit of their sum is correctly represented, whether that sum is greater than or less than 10. The coding for the 10 decimal digits is set out in Table 1.

Table 1

BINARY CODING FOR RESIDUE REPRESENTATION

Decimal digit	Residue modulo 5	Residue modulo 2	Binary coding
0	0	0	0000
1	1	1	0011
2	2	0	0100
3	3	1	0111
4	4	0	1000
5	0	1	0001
6	1	0	0010
7	2	1	0101
8	3	0	0110
9	4	1	1001

Two addition tables are necessary to carry out addition, one for each of the moduli. Their logic is shown in Table 2.

Table 2

ADDITION-TABLE LOGIC

(*a*) Modulus 5.

Second input number \ First input number	0	1	2	3	4
0	0	1	2	3	4
1	1	2	3	4	0
2	2	3	4	0	1
3	3	4	0	1	2
4	4	0	1	2	3

(*b*) Modulus 2.

Second input number \ First input number	0	1
0	0	1
1	1	0

The decimal carry has to be determined separately, and wi occur if at least two of the following three conditions are satisfie (a and b are the two input decimal digits):

(i) $a \geqslant 5$.
(ii) $b \geqslant 5$.
(iii) The sum of the residues, modulo 5, of a and b is equal to c greater than 5.

A further condition for a carry is that

(iv) The sum of a and b is equal to 9 and there is a carry fro the last decimal place.

The addition of the decimal carry consists in adding unity t both the modulus 5 and modulus 2 parts of the sum in the nex most significant decimal place.

Owing to the use of the residue system of representing th decimal digits, no interaction occurs between the two sections c the adder at any time during the addition process, including th addition of the decimal carry, or during the conversion to an from binary coding.

(5) SUBTRACTION

Subtraction is carried out by representing the number to b subtracted as a 9's complement. If the numbers are now adde together and a carry is added into the least significant decim place of the sum, the result will be the difference expressed i 10's complement form.

Examination of Table 1 shows that the 9's complement of decimal digit is obtained by permuting the digits in both th modulus 5 and modulus 2 sections, as shown in Table 3.

Table 3

COMPLEMENTS IN THE RESIDUE SYSTEM

Residue modulo 5		Residue modulo 2	
Before complementing	After complementing	Before complementing	After complementing
0	4	0	1
1	3	1	0
2	2		
3	1		
4	0		

(6) THE COMPLETE DECIMAL ADDER

The operation of decimal addition based upon the principle just described will now be summarized.

Referring to Fig. 1, each decimal digit is received as fou binary digits in parallel. It is first decoded so that one puls appears on each of two groups of output leads. The first grou containing five leads represents the residue modulo 5 and th second group of two leads represents the residue modulo 2. I addition, one of the input numbers passes through a comple menter which, if necessary, effects the permutation shown i Table 3. The decoded versions of a and b then pass to tw addition tables which are made up of coincidence gates in th arrangement shown in Table 2. Finally the sum is re-code into binary form, the digits in the modulus 5 and modulus channels having been advanced by unity if there is a carry fro the last decimal place.

The conditions under which a decimal carry will occur hav already been mentioned. To find out if the first two condition are satisfied, the inputs to the addition tables are examined. Th third condition is obtained from the modulus 5 addition tabl If the inputs are such that their sum is below and to the right c the line of 4's in Table 2(*a*), then the sum of the residues is equa

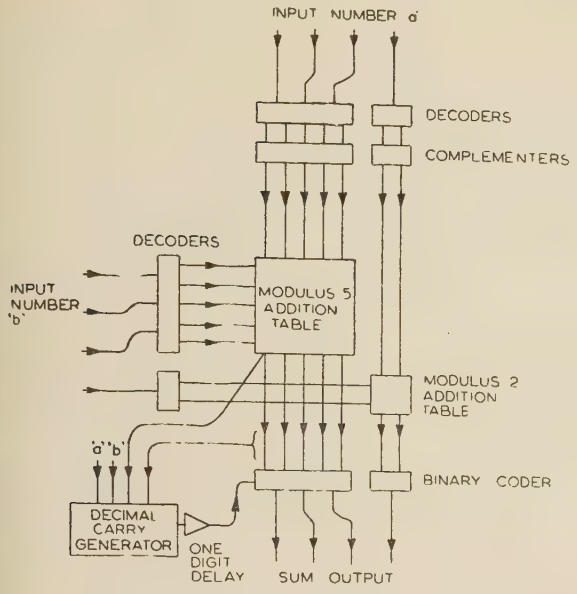


Fig. 1.—Block diagram of decimal adder.

or greater than 5. The appropriate outputs of the addition tables are examined, together with the carry from the last decimal place, to see if condition (iv) for a carry is satisfied.

(7) THE CORE-TRANSISTOR CIRCUIT ELEMENT

The core-transistor circuit element has been described in References 2 and 3. It is now proposed to discuss its design in more detail.

As shown in Fig. 2, the square-loop magnetic core has an output

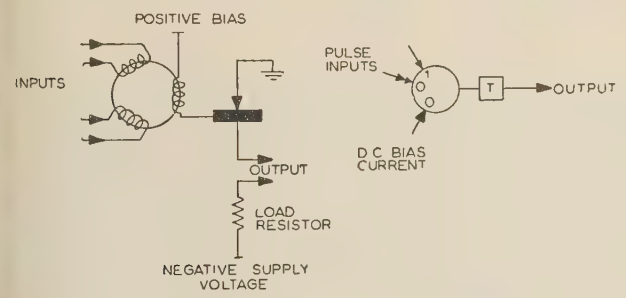


Fig. 2.—Basic core-transistor circuit.

winding which is connected in series with the base circuit of a common-emitter junction transistor. Also in series with the base is a d.c. bias voltage which maintains the transistor in a non-conducting state when there is no flux change in the core. When the flux in the core is changing in one of its two possible directions the voltage across the output winding will take the base negative, overcoming the bias, and the transistor will be switched on. Collector current will then flow and, if the base current is sufficient, the transistor will bottom and the current in the collector circuit will be limited by the load resistance. After the core has changed its state, reverse base current flows from the bias source and the transistor is switched off.

To make a satisfactory logical element, the circuit needs to emit pulses which are standardized in both amplitude and duration. The maximum amplitude of the pulses is defined by the load resistance, and they can be prevented from becoming too small by limiting the number of further cores driven by the stage under consideration.

The overall pulse length is the sum of the switching time of the core and the time taken for the transistor to switch off after the core has ceased to drive it. The switch-off time is dependent on the base bias voltage, and varies from one transistor to another. Factors which affect the switching time of the core are the input characteristics of the transistor and the amplitude of the driving pulse. This is expressed by the following equations:[3]

$$(F - N_0 I_b - k)\Delta t \simeq S \quad . \quad . \quad . \quad . \quad (1)$$

$$V_b + V_{bias} \simeq \frac{N_0 \Delta \Phi}{\Delta t} \quad . \quad . \quad . \quad . \quad (2)$$

It is found that the base voltage of a bottomed transistor is fairly constant over a wide range of base currents, as shown in Fig. 3, the two transistors whose input impedances are plotted

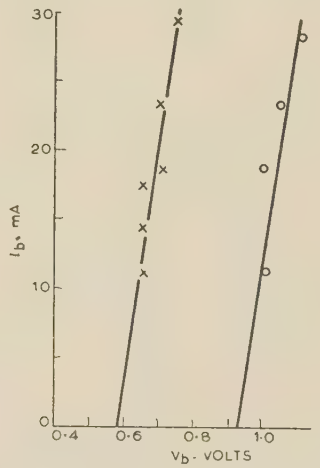


Fig. 3.—Transistor base impedance.
$I_c = 55$ mA, $V_c \simeq 0$

being at each end of the spread of voltages. These results can be represented approximately by

$$I_b = c(V_b - d) \quad . \quad . \quad . \quad . \quad . \quad . \quad (3)$$

where c and d are constants. The parameter c was found to be fairly constant for the transistors which were tested, while d had a spread of nearly 0·4 volt.

Eqns. (1), (2) and (3) result in expressions for Δt and I_b which

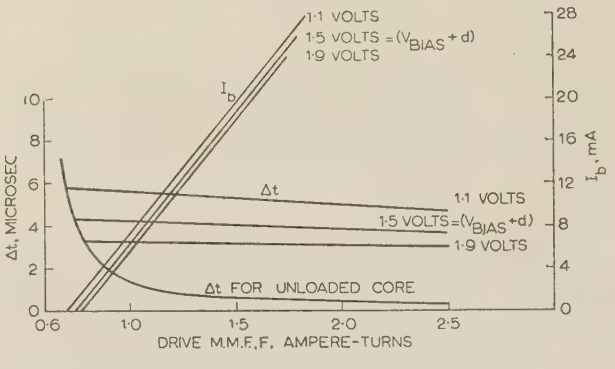


Fig. 4.—Core switching time and output current.

are plotted in Fig. 4 with the following values of the various parameters:

$S = 0{\cdot}6 \times 10^{-6}$ ampere-turn-sec.
$N_0 = 40$ turns.
$k = 0{\cdot}6$ ampere-turn.
$\Delta\Phi = 0{\cdot}16 \times 10^{-6}$ weber.
$c = 0{\cdot}18$ mho.

The value of c corresponds to an OC72 transistor and the values of S, k and $\Delta\phi$ to two FX1508 matrix storage cores treated as a single core. Experiment has shown good agreement with these curves, except for low base currents where the simple linear law assumed for the transistor input characteristic no longer holds.

It can be seen that a high value of V_{bias} is desirable to minimize the effect of variations in d between one transistor and another as well as the effect of changes in drive amplitude. Another argument in favour of a high bias voltage is that it speeds the switching off of the transistor and therefore reduces the variations in pulse duration from this source. On the other hand, the drive requirements for the circuit are lowered by a small value of V_{bias}, which allows a smaller N_0 to be used for a given pulse duration.

The amount of base current necessary to switch the transistor on in a given time is a function of its cut-off frequency and the collector current. The poor frequency response of available transistors calls for a large base current to obtain a rise time of the order of one microsecond, although a large variation has been found between different transistors of the same type.

The final design of the core-transistor element is a compromise between the various conflicting factors. The base bias voltage should be made no greater than is necessary to obtain a satisfactory control of pulse duration, in order to keep the drive requirements as low as possible. Once the pulse duration and the bias voltage have been decided upon, N_0 can be found from eqn. (2) and the transistor input-impedance data.

The number of turns on the input windings should be as low as possible for two reasons. First, it increases the number of input windings which can be provided and therefore the logical flexibility of the circuit. Secondly, since the collector supply voltage cannot be increased beyond a certain limit, a greater number of load cores can be connected in series with the collector circuit if they have smaller input windings. This process cannot be carried too far, since the large collector current which becomes necessary, and which, in turn, demands an increased I_b and F, exceeds the capabilities of the transistor.

The circuit elements actually used in the construction of the adder have the following circuit constants:

Magnetic core: two FX1508 storage cores.
All core windings: 40 turns of 48 s.w.g. wire.
Transistor: OC72.
Base bias voltage: +0·5 volt.
Base current: 24 mA.
Collector supply voltage: −20 volts.
Collector load resistance: 470 ohms.
Collector current: 42·5 mA.

The chosen value of V_{bias} limits the combined spread in pulse duration due to changes in the parameter d and the switch-off time to ± 1 microsecond. A batch of 50 transistors tested in this circuit gave overall pulse lengths varying from 6·0 to 8·0 microsec. In all cases the current rise-time was 1·2 microsec or better.

It was found possible to put six separate windings on each core, thus allowing five inputs. Each core-transistor element was capable of providing satisfactory drive to about four others before the rise-time of the driven stages became too long.

(8) CORE-TRANSISTOR LOGICAL CIRCUITS

To perform logical operations a magnetic core must behave as an analogue discriminator, able to distinguish between different drive amplitudes. In the magnetic matrix store the square hysteresis loop alone is used for this purpose. However, as has been pointed out elsewhere,[4] the use of a d.c. bias passed through a separate winding on the core to provide a threshold offers advantages. Because of the improved discrimination which bias provides, the signal pulses need not have such accurately defined amplitudes and a greater range of logical operations is possible. In addition, the bias automatically resets the cores to a standard state after the signal pulses have ceased, making possible two classes of logical elements which emit their output pulses at different times.

It is assumed hereafter that all signal pulses and bias currents have a standard amplitude, this being equal to the collector current of a stage which is emitting an output pulse and such that, when passed through a standard input winding, a drive m.m.f. equal to F (Section 7) is produced.

Fig. 2 shows a symbolic method of depicting a logical circuit element. The solid arrow pointing towards the core represents a standard winding carrying a direct current bias. Other arrows represent pulse inputs, a double arrow signifying that two standard windings are connected in series. In all cases the figure next to the arrow head indicates the direction in which the input current is driving the core. It is assumed that the transistor is switched on when the core is changing from the '1' to the '0' state.

As an example of a logical element, consider the coincidence gate shown in Fig. 5. Two simultaneous pulse inputs are necessary to overcome the bias and provide a net drive of F to the core. When the pulses cease the core reverts to its initial state, again with a net drive equal to F. Note that two changes of state are involved in this action and either one can be used to provide an output by suitably polarizing the output winding. Thus the gate can take two forms, one providing an output straight away which will be referred to as an A-period gate, and the other giving a delayed output which will be called a B-period gate. The waveforms associated with these two types of gate are shown in Fig. 5. By providing two bias windings, a triple coincidence can be detected with the same basic arrangement, and the possibilities are limited only by the number of windings available.

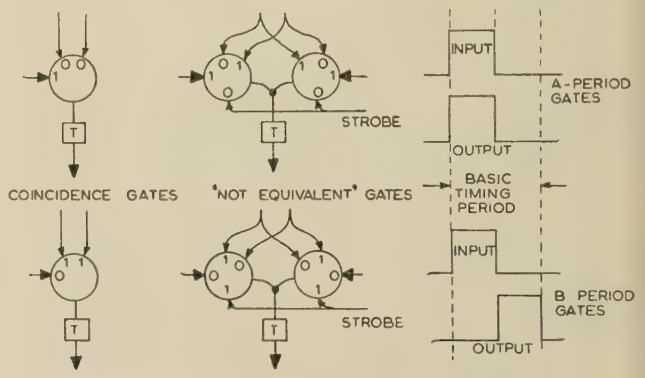


Fig. 5.—Core-transistor logical circuits.

Operations which involve the inhibiting of one pulse by another need a slightly different treatment. Because the signal pulses cannot be relied upon to have precisely equal durations, there exists the possibility of spurious operation. This difficulty is overcome by using a strobe pulse which starts slightly later than the signal pulses and finishes slightly earlier. The core can therefore change state only in the strobing period, during which time the signal inputs, if present, are at full amplitude. The 'not equivalent' gates in Fig. 5 are an example of this type of operation. Two cores, with their output windings connected in series with the base circuit of one transistor, are used to detect the two conditions which will give rise to an output. Like the coincidence gate, the 'not equivalent' gate can take A- and B-period forms.

Since each core undergoes two changes of state in performing a logical operation, the period of the timing cycle is made at least equal to the length of two signal pulses. In practice, owing to the delay of about one microsecond before an energized

element emits an output pulse, some extra time must be allowed, and a timing period of 20 microsec has been found satisfactory.

In designing a complex logical network, the operations between input and output are divided into a number of short chains. Each chain begins and ends with storage cores, i.e. cores which are not supplied with direct current bias and which only serve to store the signal temporarily. By this means the switching delays in the logical elements are not allowed to accumulate and thereby interfere with the timing. During the A part of the timing period the storage cores are 'read' and the signal then passes through one or more of the A-period type of gate. The last element in the chain is a B-period gate, which accepts the outputs of the preceding gates during the A-period and emits an output during the B-period to the storage core which terminates the chain. By this time the storage core is ready to receive new information, having disposed of that which it contained at the beginning of the timing period. If an element which involves inhibiting has to be used, it should be made the first in the chain in order to maintain the strobing tolerances.

The foregoing principles have been used in the design of the logic for the decimal adder and further details will now be given.

(9) THE DECODERS

The decoders for the modulus 5 and modulus 2 channels are entirely separate. The one for the modulus 5 channel is shown in Fig. 6.

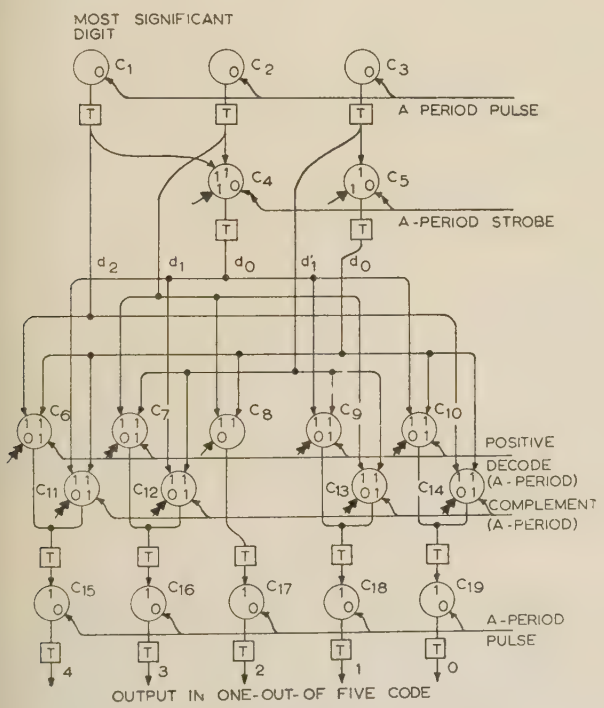


Fig. 6.—Modulus 5 decoder and complementer.

The input is three binary digits in parallel and output is required in the form of a pulse on one of five output leads corresponding to the five possible values of the input. The decoder comprises a single chain of logical operations, cores C_1–C_3 and cores C_{15}–C_{19} being the input and output storage cores, respectively. It is convenient to regard the three input digits as being split into two groups, as shown in Table 4. The first group, comprising the two most significant digits, can take three values, 0, 1 and 2. Using core C_4 to detect the case where both the digits are zero, one of the leads d_0, d_1 or d_2 will be energized. In a similar manner, and with the help of C_5, the least significant digit of the input causes one of the leads d_0 or d_1 to be energized. The final outputs are now obtained by coincidences between one of the group d_0–d_2 and one of d_0' or d_1'.

Table 4

THE LOGIC OF THE MODULUS 5 DECODER

Input digits	d_0	d_1	d_2	d_0'	d_1'	Output positive	Output complemented
000	1	0	0	1	0	0	4
001	1	0	0	0	1	1	3
010	0	1	0	1	0	2	2
011	0	1	0	0	1	3	1
100	0	0	1	1	0	4	0

Cores C_6, C_7, C_9 and C_{10} are B-period coincidence gates which require three inputs simultaneously present in the A-period to produce an output. Two of the inputs are obtained in the manner just described, and the third is a controlling pulse which only activates the cores when the number is required in positive form. When the number is required to be complemented, cores C_{11}–C_{14} are activated and take the place of cores C_6–C_{10}. It should be noted that the same combination of input digits corresponds to the output labelled '2' whether the input is being complemented or not, so that only one core is needed to produce this output, namely C_8, and this can be a 2-input gate.

The output of cores C_6–C_{14} occurs in the B-period and is stored in one of the cores C_{15}–C_{19}, from which it passes to the addition table and the carry generator during the next timing period.

The decoder for the modulus 2 channel is similar to the one just described. Those for the other input number are the same except that there need be no provision for complementing.

(10) THE ADDITION TABLES

The addition tables are arrays of A-period coincidence gates which are grouped according to the outputs they give. A typical group in the modulus 5 addition table is shown in Fig. 7, and

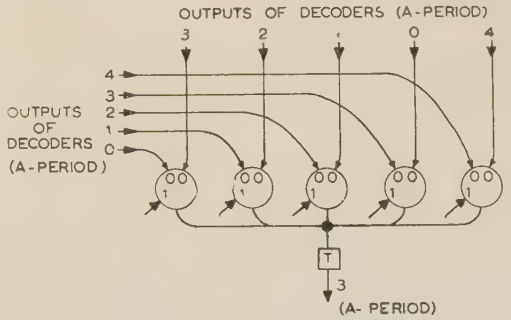


Fig. 7.—Typical addition-table element.

since there are five possible sum outputs, there are five such groups. One further output is required for generating the decimal carry and this is provided by two groups of cores similar to those shown in Fig. 7 with the collectors of the two transistors connected together. The inputs to the cores in this case are such that they correspond to elements lying below and to the right of the line of 4's in Table 2(*a*).

(11) THE CODING UNIT

The coding unit is an array of B-period gates which accepts the outputs of the addition tables during the A-period and delivers

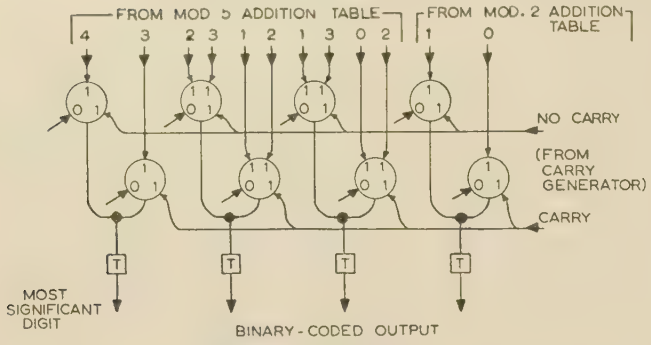


Fig. 8.—Coding circuit.

a binary-coded output during the B-period. If no decimal carry is present the top row of cores shown in Fig. 8 is activated, while in the presence of a carry the bottom row of cores advances the numbers in the modulus 5 and modulus 2 channels by unity.

(12) THE DECIMAL-CARRY GENERATOR

Fig. 9 shows the circuit for generating the decimal carry. Cores C_1 and C_2 determine whether the input number *a* is equal

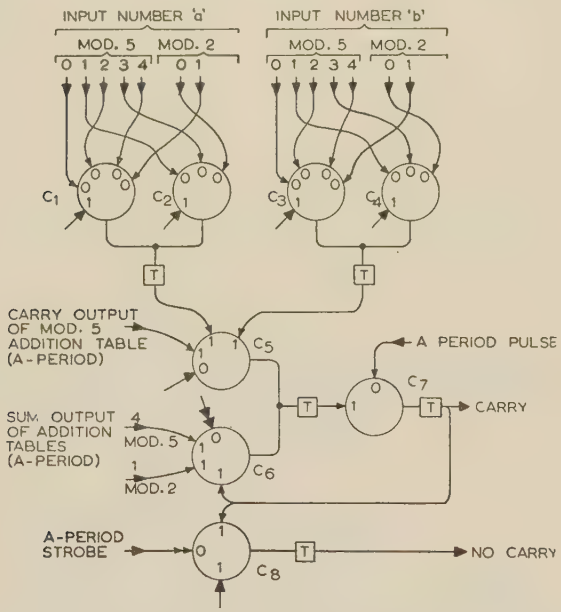


Fig. 9.—Decimal-carry generator.

to or greater than 5 by examining the appropriate digits in the modulus 5 and modulus 2 channels. Cores C_3 and C_4 perform the same function for the other input number, *b*.

The outputs of these circuits are applied to core C_5, together with the carry output from the modulus 5 addition table. At least two of these three conditions have to be satisfied for a carry to be generated and, if they are, C_5 causes a carry to be stored in C_7 during the B-period. Core C_6 generates a carry when the output of the addition tables is 9 and there is a carry from the last decimal place.

If, at the beginning of any basic period, there is no carry stored in C_7, core C_8 generates a 'no carry' pulse which is sent to the coding unit.

(13) CONTROL PULSES

Although the addition of two particular decimal digits takes two basic timing periods, digits are fed in at the full repetition rate. Thus, when a particular digit is entering the inputs of the decoders, the previous one is being stored preparatory to entering the addition tables.

The pulses which read the contents of the storage cores and activate the various gates occur during every timing period and are generated by core-transistor circuits. Of the two pulses which operate the complementer in Fig. 6, only one is required, depending on whether addition or subtraction is being carried out.

(14) TESTING METHODS

In the absence of equipment for storing complete decimal numbers, the adder has been tested only with a limited type of input. Two registers were provided, each capable of storing one decimal digit, and were connected to the two inputs of the adder. The first of these had provision for manual loading and could be arranged either to clear itself after it had been read once or to retain its contents indefinitely. The second, regarded as an accumulator, received its input from the output of the adder and had no inherent regeneration facility.

With these arrangements it was possible to transfer single digits from the manually loaded register to the accumulator, and then to pass them repeatedly through the adder. Alternatively, using the regeneration facility of the manually loaded register, the contents of the accumulator could be altered on every passage through the adder, giving rise to a repetitive pattern of digits.

Tests showed that the operation of the adder was unaffected by the repetition frequency used, provided that the basic timing period was made at least 15 microsec. It is considered that a timing period of 20 microsec provides ample margins of safety.

Satisfactory operation was obtained when the bias current and collector voltage supplies were independently varied within 5% limits in either direction. If these two supplies varied together, it was found that the tolerances were increased to about 10%. This is to be expected, since in both the A- and B-period gates the variations in drive pulse amplitudes and bias current cancel each other, provided that they occur in unison. Since it is possible to build a voltage stabilizer using transistors which will supply large currents with less than 5% variation, it is not expected that these tolerances will cause any difficulty in practice.

About 80 transistors and slightly more than twice that number of magnetic cores were used in the construction of the adder, its registers and the pulse-generation equipment. The only components which have failed are several of the cores, which developed open-circuited windings shortly after installation. Those which survived the initial period of operation have performed satisfactorily thereafter.

(15) CONCLUSION

The successful operation of the adder has shown that the core-transistor circuit has sufficient versatility and flexibility to enable it to be used in computing equipment operating at medium speeds. The component limiting the speed at present is the transistor, and it is to be expected that improved transistors will greatly extend the scope of the circuit when they are available.

The fact that all signals and waveforms associated with this type of circuit are pulses necessitates a slightly different approach to the problems of logical design. The use of biased cores, leading to the possibility of A- and B-period gates, has enabled logical operations to be performed economically without loss of speed.

Although the logic of the adder was designed with the particular properties of the core-transistor circuit in mind, an adder using this type of number representation could equally well be realized with the more conventional circuit techniques. In this case the temporary storage between the outputs of the decoders

and the inputs of the addition tables would probably be discarded, together with the A- and B-period type of logic. The delay before an output was obtained would then be limited only by the response time of the intermediate gates. The elimination of the time required for carry propagation would still be obtained, as well as the other advantages mentioned in Sections 3 and 4. It is probable that an adder using diode matrices and valves could be operated at a speed of at least 10 times that which is possible with the transistor-core combination in its present state of development.

(16) ACKNOWLEDGMENTS

The authors would like to thank Dr. T. Kilburn and Dr. D. B. G. Edwards for their interest and encouragement. Mr. Maclean wishes to thank the New Zealand Defence Science Committee for making it possible for him to carry out this work. Mr. Aspinall acknowledges a grant from N.R.D.C.

(17) REFERENCES

(1) COURANT, R., and ROBBINS, H.: 'What is Mathematics?' (Oxford University Press, 1941).

(2) PERRY, G. H., HOFFMAN, G. R., and SHALLOW, E. W.: 'A New and Simple Type of Digital Circuit Technique using Junction Transistors and Magnetic Cores', *Proceedings I.E.E.*, Paper No. 2112 M, November, 1956 (**103** B, Suppl. 3, p. 412).

(3) HOFFMAN, G. R., and MACLEAN, M. A.: 'Quiescent Core-Transistor Counters', *ibid.*, Paper No. 2122 M, November, 1956 (**103** B, Suppl. 3, p. 418).

(4) SCARROTT, G. G., HARWOOD, W. J., and JOHNSON, K. C.: 'The Design and Use of Logical Devices using Saturable Magnetic Cores', *ibid.*, Paper No. 2065 M, March, 1956 (**103** B, Suppl. 2, p. 302).

(18) APPENDIX

Let a positive integer a be divided in turn by positive integers $m_1, m_2, \ldots m_s, \ldots m_n$, leaving remainders $r_1, r_2, \ldots r_s, \ldots r_n$.

Then $$a = l_s m_s + r_s \text{ for all } s$$

where $r_s < m_s$ and l_s is a positive integer.

If there exist two integers a_1 and a_2 which give the same set of remainders,

then $$a_1 = l_{1s} m_s + r_s$$

and $$a_2 = l_{2s} m_s + r_s \text{ for all } s$$

Taking a_1 to be greater than a_2, it then follows that $a_1 - a_2$ is a positive integral multiple of m_s. Since this is true for all values of s from 1 to n, it follows that

$$a_1 - a_2 = LM$$

where L is a positive integer and M is the l.c.m. of $m_1, m_2, \ldots m_s, \ldots m_n$.

[The discussion on the above paper will be found on page 144.]

621.317.7.087.6 : 621.374.32 : 681.142

The Institution of Electrical Engineers
Paper No. 2441 M
Oct. 1957

AN ACCURATE ELECTROLUMINESCENT GRAPHICAL-OUTPUT UNIT FOR A DIGITAL COMPUTER

By T. KILBURN, M.A., Ph.D., D.Sc., Member, G. R. HOFFMAN, Ph.D., B.Sc., and R. E. HAYES, M.Sc., Student.

(*The paper was first received 1st August, and in revised form 3rd September, 1957. It was published in October, 1957, and was read before the* MEASUREMENT AND CONTROL SECTION *5th November, 1957.*)

SUMMARY

The paper describes a graphical-output unit fabricated from a uniform layer of electroluminescent phosphor with 512 parallel conducting strips on either side, these respective sets of strips being mutually at right angles to form a matrix of conductors. The application of a changing electric field between selected strips of both sets causes the phosphor to fluoresce at their intersection. As intersections are selected in turn a pattern can be traced out on the panel and recorded photographically. The discrimination ratio of brightness of wanted to unwanted spot positions and an improved method of biasing the matrix are considered in some detail. Other factors, such as driving-voltage frequency, waveshape and intermittent operation, are considered as they affect the discrimination ratio. Details of some prototype matrices are given, and a display indicating the resolution of such a system is shown.

LIST OF SYMBOLS

B = Brightness or integrated light output (i.l.o.) for an applied voltage V.
V = Peak-to-peak applied voltage.
a, b = Constants for a certain phosphor at a particular frequency and voltage waveshape.
B' = Brightness for an applied voltage of $\frac{1}{2}V$.
D' = Discrimination ratio—the ratio between the two levels of brightness B and B', i.e. $D' = B/B'$.
B'' = Brightness for an applied voltage of $\frac{1}{3}V$.
D'' = B/B''.
d = Thickness of electroluminescent layer.
D_1'' = Discrimination ratio similar to D'' but for electroluminescent cell thickness $\frac{1}{2}d$.
E = Electric field across the electroluminescent cell.
B_1 = Another brightness level.
D_2'' = Discrimination ratio similar to D'' but for brightness level B_1.
A = Waveform peak-to-peak amplitude $\frac{1}{2}V$ volts.
B = Waveform peak-to-peak amplitude $\frac{1}{2}V$ volts but antiphase to A.
α = Waveform peak-to-peak amplitude $\frac{1}{6}V$ volts in phase with B.
β = Waveform peak-to-peak amplitude $\frac{1}{6}V$ volts in phase with A.

(1) INTRODUCTION

The results of calculations made on a digital computer appear at the output in numerical form; in many cases, however, it is desirable to display them graphically, which may involve many hours of work. There is, therefore, a demand for an automatic device which will provide an accurate graph whenever this is required. Furthermore, as information is easier to assimilate in graphical than in tabulated form, a rapid assessment of any set of results can be made and the computer input parameters appropriately changed during the progress of a particular problem.

Graphical output can be provided at the screen of a conventional cathode-ray tube. In this case the digital-computer output is converted into analogue form to produce deflections of an electron beam along the X- and Y-axes, the display being built up sequentially point by point. The screen can be observed visually, as a complete trace of, say, one thousand spot positions can be completed in a few seconds: alternatively, the display can be recorded photographically. The accuracy of such a system is limited by the inherent errors in the analogue circuit parameters, the varying deflection sensitivity and the varying angle between X- and Y-axes over the screen surface. In addition, the short-term stability of such a system may be adequate but the long-term stability, unless elaborate circuit arrangements are used, will be poor. In the electroluminescent panel to be described here, these defects are eliminated because the 'spot' can only take positions which are at the intersections of two fixed sets of conducting strips.

The panel (Fig. 1) is made from a layer of suitable electroluminescent phosphor with a set of parallel conducting strips disposed on either side. These respective sets of strips are mutually at right angles and form a matrix of conductors.[1] The application of a changing electric field between selected strips of both sets causes the phosphor to fluoresce at their intersection. By appropriate selection of pairs of conducting strips in turn a pattern is traced out on the panel, and as one set of the conductors is transparent the pattern can be photographed through the glass support. With a system of this type, digital to analogue conversion is unnecessary, and the accuracy is now of high order, since it is dependent upon a master negative which is fabricated to fine limits. The long-term stability is, of course, inherent and, furthermore, even a replacement panel is identical.

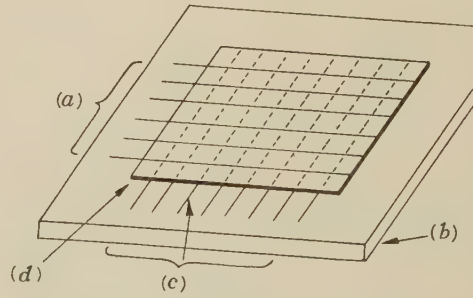


Fig. 1.—Basic graphical display panel.
(*a*) Opaque conductors.
(*b*) Glass.
(*c*) Transparent conductors.
(*d*) Phosphor in plastic.

The authors are in the Electrical Engineering Department, University of Manchester.

(2) ELECTROLUMINESCENCE

The phenomenon of electroluminescence, the sustained emission of light by a phosphor subjected to the action of an alternating electric field, was first reported by Destriau[2] in 1936. Since 1946 a large number of papers on the subject have been published, and much effort has been expended in producing phosphors for use as new light sources.[3, 4] So far, the application has been limited, as the efficiency is low and phosphor panels are not ideally suited to the domestic supply (since the brightness increases with the frequency). However, at lower light levels, e.g. dark-room illumination, exit signs, and at higher frequencies, e.g. instruction signs in aircraft, instrument-panel illumination, phosphor panels have been used to advantage. In conjunction with photoconductive elements, some interesting electronic devices have also been constructed. The brightness of the panels depends on the amplitude, waveform and frequency of the applied voltage. The spectral distribution of the emitted light depends on the phosphor used and on the frequency of excitation. The most commonly used phosphor is zinc sulphide which has been activated by the addition of impurities of copper or silver.

Simple electroluminescent cells can be made as shown in Fig. 2.

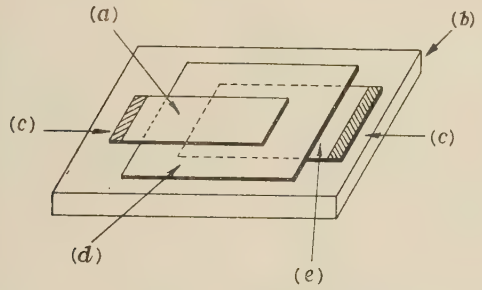


Fig. 2.—Electroluminescent cell.

(*a*) Aluminium.
(*b*) Glass.
(*c*) Indium contact.
(*d*) Phosphor.
(*e*) Transparent conducting bismuth-oxide/gold layer.

A transparent conducting layer of bismuth oxide and gold[5] is evaporated on to a glass sheet to form one electrode. On to this is sprayed an even layer of phosphor suspended in a polymethacrylic ester; the layer must be thin—not more than a few mils thick—in order to obtain sufficiently large field strengths for luminescence to be observed. Metal backing electrodes, opaque layers of aluminium or copper, are then evaporated on to the phosphor. Contacts to these electrodes are made with indium solder fused to the glass surface.

Measurements show that the average brightness of an electroluminescent cell increases rapidly with increase of applied alternating voltage. Over a wide range of brightness the dependence can be represented by[6]

$$B = a\varepsilon^{-b/\sqrt{V}} \quad . \quad . \quad . \quad . \quad . \quad . \quad (1)$$

When the brightness is measured with a photomultiplier and integrating amplifier (Section 10), a graph of integrated light output (i.l.o.) against $1/\sqrt{V}$ (Fig. 3) shows that the above relationship is experimentally verified over a brightness range as large as $10^8 : 1$.

Relationships for brightness as a function of frequency are less satisfactory and depend on waveshape and voltage; however, in the lower-frequency range, the brightness is approximately a linear function of frequency. Of more direct interest in the present application is the discrimination ratio, which is defined and discussed in Section 4.

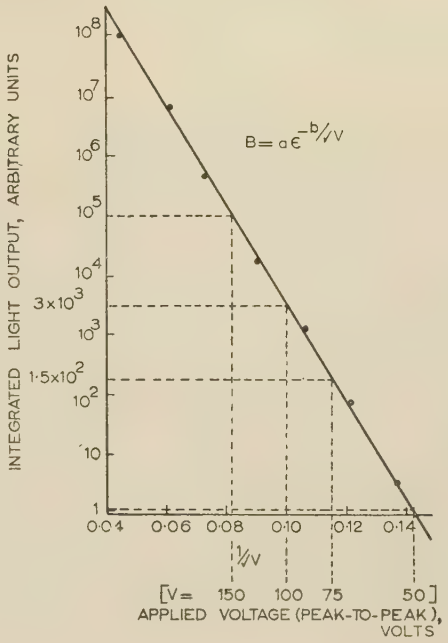


Fig. 3.—Voltage/brightness characteristic of electroluminescent cell.

(3) THE ELECTROLUMINESCENT MATRIX

The graphical-output unit will have 512 × 512 orthogonal conductors. The conductors, being 10 mils wide and spaced by 4 mils, cover an area of 7 in × 7 in. Such a panel can be regarded as being equivalent to a sheet of graph paper with more than a quarter of a million 'spot' positions. The remaining area of a 10 in × 10 in glass supporting panel is used to accommodate a 'flag' pattern of bonded silver conductors, the flags being of sufficient size to allow an electrical soldered connection to be made to each conductor. Photo-etching techniques are used throughout the manufacturing process and a complete prototype matrix with 128 × 128 conductors at 50 to the inch is more easily printed, and is shown in Fig. 4. Some of the operational

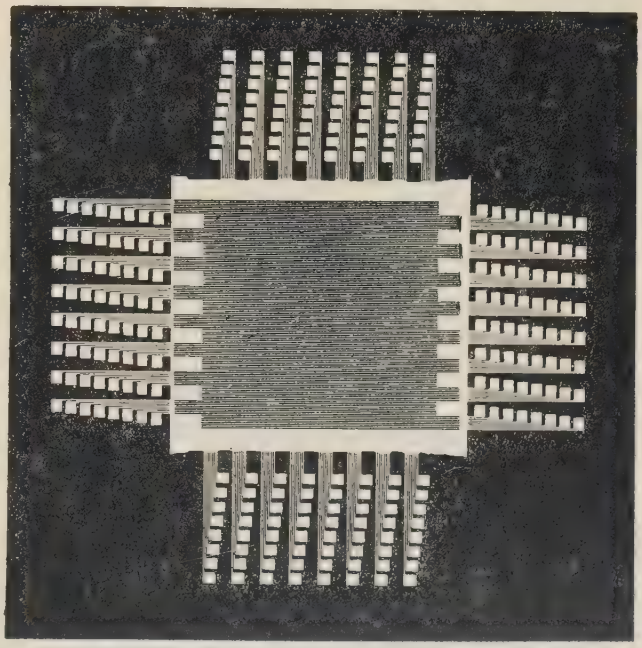


Fig. 4.—Prototype array.

features of the graphical-output unit will now be described in more detail.

(4) THE DISCRIMINATION RATIO

The electroluminescent matrix can be operated by applying voltages to the selected x- and y-co-ordinate conductors; these voltages being antiphase but having equal peak-to-peak amplitudes of $\frac{1}{2}V$ volts; a peak-to-peak voltage of V volts will then appear at the intersection x, y. If all the other electrodes are at earth potential, a peak-to-peak voltage of $\frac{1}{2}V$ appears between the selected row or column and the other earthed conductors. Unwanted spot positions along the selected ordinate and co-ordinate thus luminesce, but with lower intensity. The ratio of wanted to these unwanted brightness levels is defined as the discrimination ratio. In this graphical-output application the photographic film recording the wanted luminescent spot pattern is continuously exposed to the lower brightness levels at the unwanted spot positions. The image of these unwanted spot positions will thus accumulate as the wanted spot positions are sequentially selected. The discrimination ratio therefore gives an indication of the number of selected spot positions which may be recorded before a serious background interference becomes apparent, and this ratio must be as high as possible.

It is now advantageous to consider eqn. (1) in more detail. Since

$$B = a\varepsilon^{-b/\sqrt{V}}$$

and if

$$B' = a\varepsilon^{-b/\sqrt{\frac{1}{2}V}}$$

where B' is the brightness level for an applied voltage of $\frac{1}{2}V$, the discrimination ratio D' can be defined as the ratio of these two levels of brightness, i.e.

$$D' = \frac{B}{B'} = \varepsilon^{+0\cdot414b/\sqrt{V}} \quad . \quad . \quad . \quad . \quad . \quad (2)$$

This shows that the ratio increases rapidly as the voltage (and of course the brightness) is lowered; and, with the panels constructed so far, a discrimination ratio which allows one or two thousand spots to be recorded demands a relatively low level of brightness.

(4.1) An Improved Method of Biasing the Matrix

A considerable improvement in the discrimination ratio is obtained when the unselected conductors are subjected to an antiphase voltage waveform of one-third the amplitude of the selected conductor (Fig. 5). [The relative amplitudes and phasing of the driving waveforms A, B, α, β are shown later in Fig. 13(a)–(d).] In this case, when a voltage V appears at the intersection of the selected conductors, all the unselected intersections are subject to a peak-to-peak amplitude of $\frac{1}{3}V$ volts.

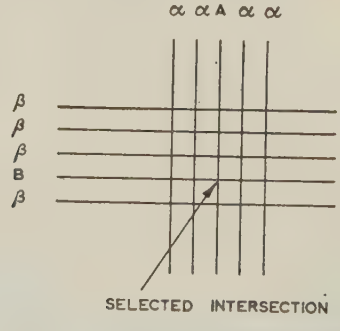


Fig. 5.—Improved biasing of matrix.

A = Waveform peak-to-peak amplitude $\frac{1}{2}V$ volts.
B = Waveform peak-to-peak amplitude $\frac{1}{2}V$ volts but antiphase to A.
α = Waveform peak-to-peak amplitude $\frac{1}{3}V$ volts in phase with B.
β = Waveform peak-to-peak amplitude $\frac{1}{3}V$ volts in phase with A.

Now if

$$B'' = a\varepsilon^{-b/\sqrt{\frac{1}{3}V}} \quad . \quad . \quad . \quad . \quad . \quad . \quad (3)$$

where B'' is the brightness level for an applied voltage of $\frac{1}{3}V$, the improved discrimination ratio, D'', of the two brightness levels corresponding to an applied voltage of V and $\frac{1}{3}V$, is

$$D'' = \frac{B}{B''} = \varepsilon^{+0\cdot732b/\sqrt{V}} \quad . \quad . \quad . \quad . \quad (4)$$

From eqns. (2) and (4),

$$\frac{D''}{D'} = \varepsilon^{+0\cdot318b/\sqrt{V}} \quad . \quad . \quad . \quad . \quad . \quad (5)$$

Thus the improvement in discrimination ratio between this system of operation and the first method is also an exponential function, and biasing becomes increasingly worth while as the brightness levels are reduced, i.e. as the voltage V is lowered. As an example, an experimental determination of the discrimination ratio under typical operating conditions can be calculated from Fig. 3. With a peak-to-peak applied waveform of 150 volts the i.l.o. is 10^5 units; for 75 volts, $1\cdot5 \times 10^2$ units; and for 50 volts, 1·2 units. This gives $D' = 600$ and $D'' = 80\,000$, an improvement of more than 100 times. Using this method of biasing the matrix, all the unwanted spot positions are then equally bright, giving the advantage of a uniform background illumination: the background of the previous system would, of course, be pattern-sensitive. In what follows, equations and arguments assume that the biasing method of operating the panel is being used.

(4.2) Thickness of the Electroluminescent Layer

The phosphor layer must be thin in order to obtain sufficiently large field strengths with reasonable applied voltage amplitudes, and uniform in thickness to give an overall even lumination. Assuming that equal electric fields across the electroluminescent cell cause equal luminescent brightness, and since b is a constant for any particular frequency and waveshape, then eqn. (4) may be interpreted for cells of thickness d or $\frac{1}{2}d$ as

$$D_1'' = D''[\varepsilon^{0\cdot732b/\sqrt{(Ed)}}]^{0\cdot414} \quad . \quad . \quad . \quad . \quad (6)$$

where D'' is the discrimination ratio of a cell of thickness d, and D' is the discrimination ratio similar to D'' but for an electroluminescent cell of thickness $\frac{1}{2}d$. Experiments to confirm this interpretation are difficult to perform accurately, but a considerable increase in the discrimination ratio has been observed in specially constructed wedge-shaped cells. Thinner cells, then, allow either brighter operation levels with the same discrimination ratio, or the advantage of lower-amplitude driving voltages which would be convenient with semi-conductor driving systems. It will be seen that the possibility of evaporation of very thin electroluminescent layers may be of great significance for future systems of this type.

(4.3) Driving-Voltage Waveform

The voltage/brightness characteristic of the electroluminescent phosphor depends upon the frequency and the waveshape of the driving voltage waveform. Sine, square and differentiated waveforms were investigated in detail over a wide range of frequencies; Fig. 6 shows some of the results obtained. From results of this type the constant b may be calculated and recorded as a function of frequency and waveshape (Fig. 7).

Although the largest value of b seems desirable, as it represents the rate of change of log (brightness) relative to reciprocal

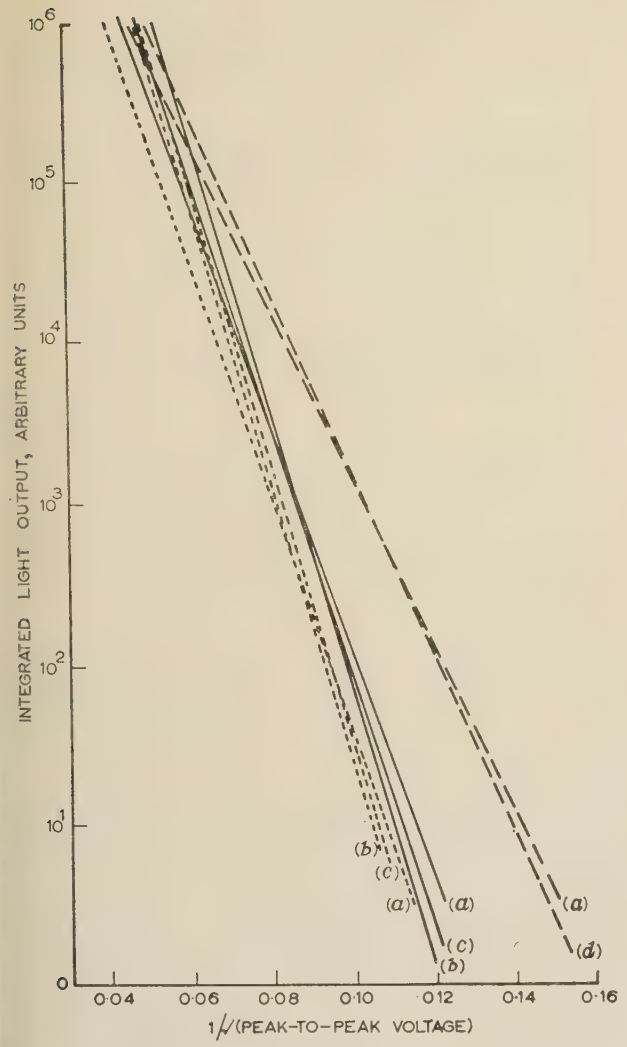


Fig. 6.—Voltage/brightness characteristics for various frequencies and waveform.

— — — — Sine-waveforms.
——— Square-waveforms.
——— —— Differentiated square-waveforms.

(*a*) 1 kc/s.
(*b*) 30 kc/s.
(*c*) 250 kc/s.
(*d*) 15 kc/s.

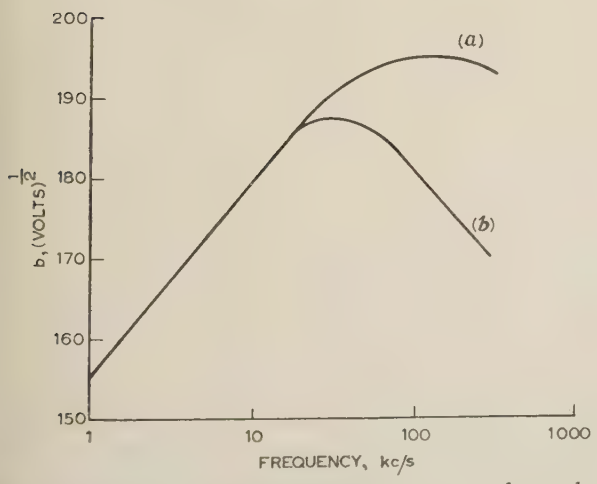


Fig. 7.—Variation of the constant *b* with frequency and waveshape.

(*a*) Sine-waveforms.
(*b*) Square-waveforms.

square root of applied voltage, this may be misleading when the highest discrimination ratio is required. The constant *a*, owing to the considerable extrapolation necessary, cannot be measured with any accuracy. More useful information from the results is obtained in the following manner.

In operation, when the graphical-output unit is controlled by a digital-computing machine, pulse information from the machine will be decoded to enable the required spot positions to be selected. Each position or graphical point must be held activated for a short period sufficient for photographic recording. Thus, in certain circumstances, the operating rate of the computer may be seriously reduced in order to match the operating rate of the panel as the points are sequentially selected and activated. It was decided that if the machine was slowed down for 30 sec to record a thousand points, the overall economic operation of the computer would not be significantly impaired. Allowing for decoding and selection time each point may then be activated for, say, 20 millisec. With this fixed exposure time the minimum phosphor brightness necessary for a photographic record with reasonable contrast can be measured. We can then obtain the peak-to-peak voltage to provide the above minimum phosphor brightness. This voltage will, of course, vary with the frequency and waveshape of the applied waveform. The brightness at a third of these voltages can now be calculated and the discrimination ratio measured as a function of frequency and waveshape

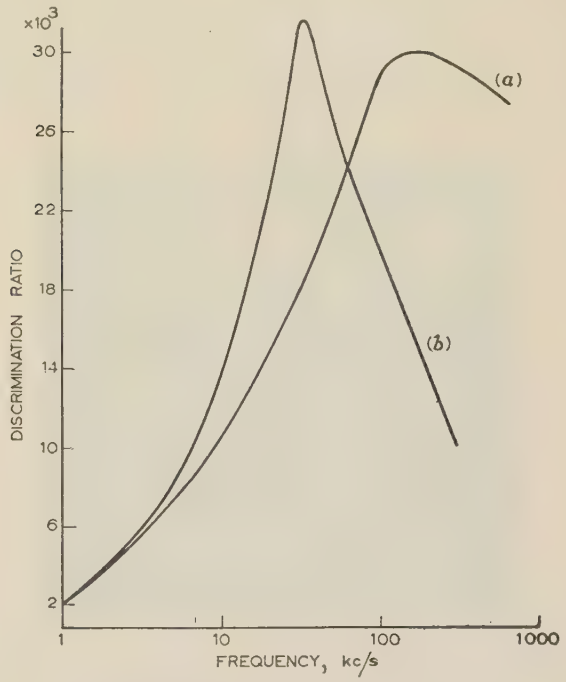


Fig. 8.—Discrimination ratio as a function of frequency and waveform.

(*a*) Sine-waveforms.
(*b*) Square-waveforms.

(Fig. 8). This gives the interesting result that the discrimination is a maximum at 30 kc/s for the square-waveform voltage drive, and a maximum at 100 kc/s for the sine-waveform drive.

(4.4) Exposure Time and Photographic Recording Material

The total number of points which may be recorded before interference from unselected points is apparent depends on the discrimination ratio. In some applications it may be desirable to record a large number of points, in which case the brightness will have to be reduced to improve the discrimination ratio, and

the exposure time accordingly increased. This may be considered in more detail as follows:

From eqns. (1) and (3) it will be seen that

$$\frac{B''}{a} = \left(\frac{B}{a}\right)^{\sqrt{3}} \quad . \quad . \quad . \quad . \quad . \quad . \quad (7)$$

At another brightness level, B_1,

$$\frac{B_1''}{a} = \left(\frac{B_1}{a}\right)^{\sqrt{3}} \quad . \quad . \quad . \quad . \quad . \quad . \quad (8)$$

The new discrimination ratio D_2'' is thus

$$D_2'' = D''\left(\frac{B}{B_1}\right)^{0\cdot732} \quad . \quad . \quad . \quad . \quad . \quad (9)$$

Therefore, if any brightness level B is halved, i.e. $B_1 = \frac{1}{2}B$, then, at the expense of doubling the time of photographic exposure, the improvement of the discrimination, D_2''/D'', is a factor of 1·66 : 1.

Conversely, the exposure time may be reduced and the speed of output raised if fewer points are to be recorded and the lower discrimination ratio can be tolerated.

The choice of photographic material is essentially a compromise between maximum contrast and highest speed, together with the consideration of blue-green sensitivity. Operating near the knee of the density/exposure characteristic of the material gives a further increase (in practice, 10 times) in the overall discrimination. To confirm this for an actual system, an early experimental 8 × 8 array with manual selection was used to record a monograph (Fig. 9). The selected spots were illuminated for 20 millisec, whereas the unwanted spots, which all had a one-third amplitude voltage across them, were illuminated for 10 min, which caused the background illumination to be only just apparent in the negative (i.e. the overall discrimination in this case was better than the calculated 1600 : 1).

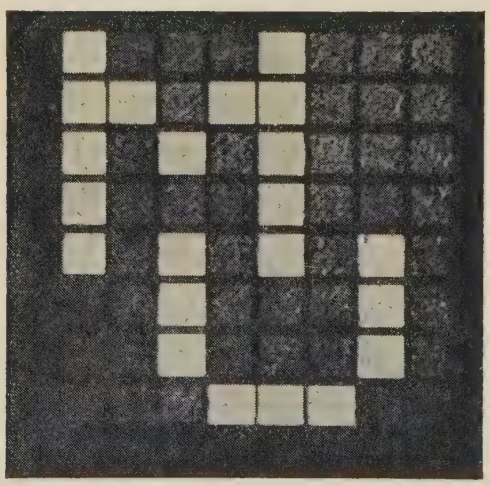


Fig. 9.—Electroluminescent monograph recorded by automatic mechanical selection.

(4.5) Simulation of the Operating Conditions of the Electroluminescent Matrix

The largest discrimination ratio for square-wave voltage drive occurs at 30 kc/s (Section 4.4). As square waves are easy to generate at this frequency and the gating and timing is a well-established computer technique, the following experiments were performed to simulate actual operating conditions of the electroluminescent matrix.

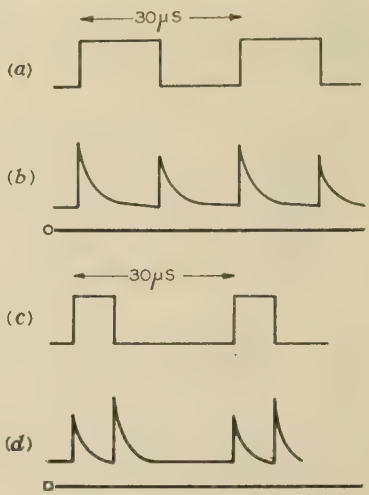


Fig. 10.—Light output pulses.

(*a*) Symmetrical applied voltage drive waveform of 150-volt amplitude and 30 microsec period.
(*b*) Light output due to (*a*).
(*c*) Asymmetrical applied voltage drive waveform.
(*d*) Light output due to (*c*).

The light output measured by a photomultiplier and amplifier when a square wave [Fig. 10(*a*)] is applied is similar to a series of unidirectional 'differentiated' pulses [Fig. 10(*b*)]. After a short delay of 0·5 microsec, the phosphor reaches its maximum brightness in 2 microsec, and the first rapid part of the hyperbolic decay has an apparent time-constant of about 10 microsec. Alternate pulses are of variable amplitude. This is caused by the reflection coefficient of the opaque electrode and the transmission coefficient of the dielectric and embedded phosphor, as the phosphor particles themselves luminesce on alternate faces depending on the direction of the applied field.[7–10] Variation of the symmetry of the applied waveform [Fig. 10(*c*)] also varies the relative amplitudes of the light output pulses [Fig. 10(*d*)]. Of more immediate concern is the variation of integrated light output as the symmetry changes. The i.l.o. is a maximum when the waveform is symmetric; however, even for a change as large as 6 : 4 in the ratio of pulse duration to pulse spacing, the i.l.o. falls by only 2% and consequently this effect may be ignored in a practical system.

(4.5.1) Build-up Time of the Electroluminescent Cell.

In operational use, any particular spot will be selected infrequently. It is thus necessary to know how the i.l.o. varies with the interval which has elapsed since it was last selected. An electroluminescent cell continuously excited with a 30 kc/s square-wave drive [Fig. 11(*a*)] gives light output pulses of the form shown in Fig. 11(*b*). (It is to be noticed that the light output does not fall to zero between pulses.) If the cell is left for a considerable time (15 hours) [see Fig. 11(*c*)] completely different light output pulses [Fig. 11(*d*)] occur during the initial 50 pulses, until the equilibrium light output [Fig. 11(*b*)] is obtained. A light output pulse is absent on the first rising edge. It is the build-up of the light pulses, coincident with the rising edges of the applied voltage, which reduces the total integrated light output for a complete 20 millisec burst of 30 kc/s square waves. This effect is compatible with the experiments performed by Waymouth[11] at low frequencies using ballistic-galvanometer techniques, and other workers using higher frequencies.[9, 12] The time (15 hours) was progressively reduced between the 20 millisec bursts of 30 kc/s square waves, and it was found that

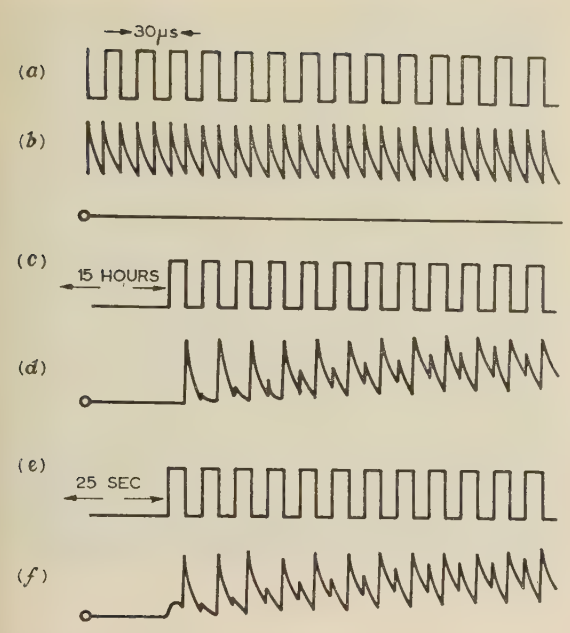
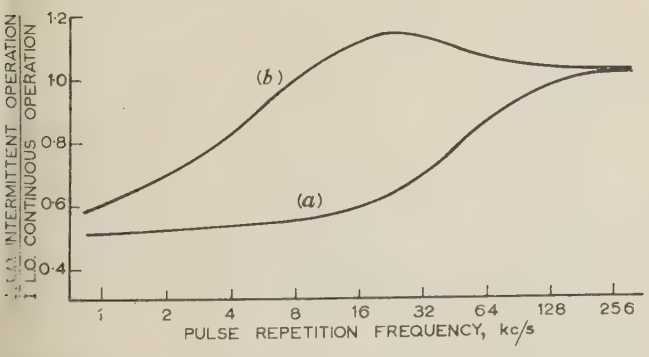


Fig. 11.—Build-up time of cell.

(*a*) Continuous voltage drive waveform of 50-volt amplitude and 30 microsec period.
(*b*) Light output due to (*a*).
(*c*) Intermittent voltage drive waveform with 15 hour rest period.
(*d*) Light output due to (*c*).
(*e*) Intermittent voltage drive waveform with 25 sec rest period.
(*f*) Light output due to (*e*).

o significant change from Fig. 11(*d*) occurred for rest periods reater than 8 min; this time can be reduced still further if an nfra-red source is allowed to 'illuminate' the cell. For periods ess than 8 min, a light pulse coincident with the first edge is pparent and increases in amplitude as this time is reduced. A typical oscillogram when the rest period is only 25 sec is hown in Fig. 11(*f*).

In the present application the integrated light output during 20 millisec period is of importance during these intermittent perating conditions. The effect of the transient conditions ecomes more apparent as the pulse repetition frequency of the pplied waveform is decreased. This is to be expected since the umber of pulses during the build-up period becomes an appreiable fraction of the total number in any 20 millisec period. In act, the i.l.o. falls to half the 'continuous' value at 5 kc/s. The atio of this transient i.l.o. to the continuous i.l.o. can be plotted or various pulse repetition frequencies and waveshapes (Fig. 12), nd the curve (*a*) of this ratio for a typical selected intersection



ig 12.—Integrated light output: ratio of intermittent operation to continuous operation for various pulse repetition frequencies.

(*a*) Typical selected intersection voltage V.
(*b*) Typical intersection voltage $\frac{1}{3}V$ at an unselected position.

voltage V approaches unity asymptotically as the p.r.f. is increased. It will be seen that for an intermittent p.r.f. of, say, 5 kc/s the light output for voltage V is only 50% of the continuous value, whereas for a voltage $\frac{1}{3}V$, curve (*b*), the intermittent light output is almost equal to the continuous value. Thus, the previously calculated discrimination ratios in this case are reduced by 50%. In order to prevent this reduction of discrimination ratio, a high pulse repetition frequency is therefore desirable.

(4.5.2) Waveforms during Change-over.

Electroluminescent cells with the phosphor suspended in insulating dielectric are sensitive to variations in electric field only; the superposition of a steady field does not affect the brightness/voltage characteristic. In Fig. 13 the waveforms (*a*),

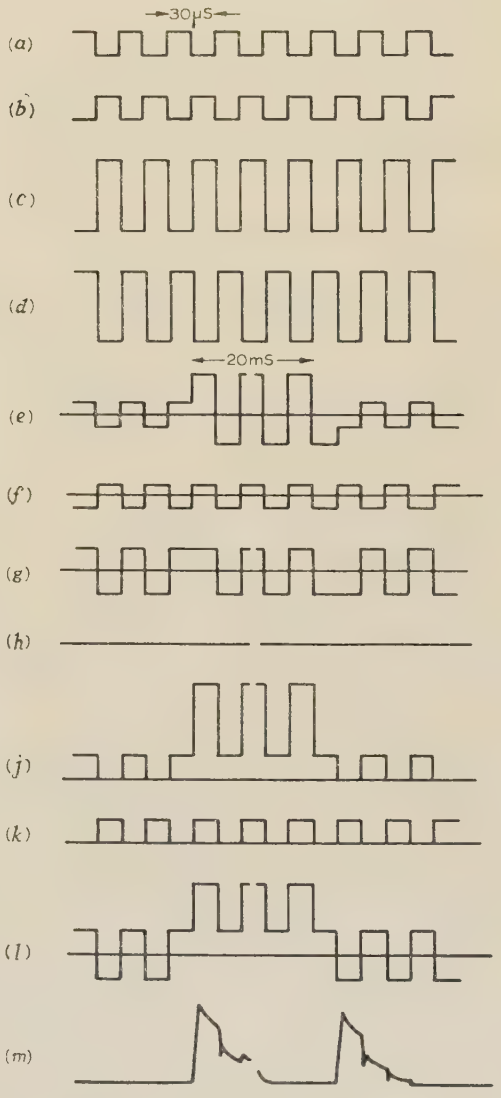


Fig. 13.—Waveforms during change-over.

(*a*) α-waveform.
(*b*) β-waveform.
(*c*) A-waveform.
(*d*) B-waveform.
(*e*) Drive waveform along selected co-ordinate.
(*f*) β-waveform.
(*g*) Electric field across cell due to (*e*) and (*f*) waveforms.
(*h*) Light output due to (*g*).
(*j*) Drive waveform along selected co-ordinate.
(*k*) β-waveform.
(*l*) Electric field across cell due to (*j*) and (*k*) waveforms.
(*m*) Light output due to (*l*).

(*b*), (*c*) and (*d*) represent the amplitude and phasing of the waveforms α, β, A and B (Fig. 5); no d.c. components have been recorded. Confining our attention to an unselected spot lying on a selected co-ordinate, the waveforms [Figs. 13(*e*) and 13(*f*)] satisfy the V to $\frac{1}{3}V$ method of operating the matrix and are shown before, during and after a change-over period. The applied electric field across the unselected spot due to (*e*) and (*f*) can be seen in (*g*).

If, however, driving waveforms of the type (*j*) and (*k*) are used which still satisfy the V to $\frac{1}{3}V$ method of operation, the field across the cell is shown by (*l*); but a field change of this type causes a series of decaying light output pulses shown by (*m*) at each transient change-over. These transients which occur along each selected co-ordinate would reduce the discrimination ratio for the whole panel. It is thus essential to arrange a symmetrical change-over, as illustrated in types (*e*), (*f*) and (*g*), since no transient light output pulses then occur [see curve (*h*)]. Some of the panels, in particular the very thin ones with poor insulation resistance, exhibit a small amount of d.c. electroluminescence. It is thus desirable to ensure that the d.c. component of the electric field is also a minimum if the brightness/alternating-voltage characteristic is to remain unmodified; further, the risk of cell breakdown is reduced.

(5) *X* AND *Y* CO-ORDINATE SELECTION

The selection of the 512*X* and the 512*Y* co-ordinates presents a formidable problem. More elegant schemes than the one to be described are being pursued, and it is hoped to publish these at a later date. In the present scheme the selection equipment is divided into two parts, relays being used for low-speed selection of 16 groups of 32 conductors, and individuals in each group of 32 conductors being selected by high-speed electronic circuits. As the graphs recorded will usually be continuous curves, the longer change-over time, which will occur relatively infrequently, will not seriously hold up the computer. Facilities for 'drawing' axes quickly may be arranged by non-selection of one set of co-ordinates and supplying a driving waveform to all the co-ordinates instead of the bias waveform. A normal selection of one of the other co-ordinates will cause 512 spots along a line to be simultaneously recorded. It is also easy to arrange for scaling markers along the axis to be rapidly recorded.

(6) THE PROTOTYPE MATRIX

After the initial experiments with 8 × 8 arrays of 1 mm square 'spots', prototype units with 128 × 128, i.e. 16 384 spots at 50 to the inch were constructed. The spot size is slightly larger than in the final 512 × 512 panel, in which the conductors are spaced 72 to the inch, but all the production techniques developed on the smaller project were directly applicable to the larger panel. A complete 128 × 128 panel is shown in Fig. 4. With this method of construction a single sheet of glass supports both sets of conducting flags for the conductors on either side of the phosphor. As an example of the resolution of the prototype, eight groups of eight conductors on both axes have been selected simultaneously and the luminescent pattern photographed, giving 64 spots at each major intersection (Fig. 14).

The final array is in an advanced state of construction; an accurate master negative has been made, and a larger evaporation plant with a rotatable substrate holder is now in operation. It is hoped that the completed unit will be installed in late 1957, soon after the Mark II Mercury computer is delivered to the Department.

(7) CONCLUSION

It has been demonstrated that an electroluminescent panel with co-ordinate selection of more than 250 points to the square inch is a practical proposition. A discrimination ratio of brightness levels between wanted and unwanted spot positions, due to the co-ordinate method of selection, has been found to depend upon the pulse repetition frequency, waveshape and intermittency of operation of the driving voltage waveforms. This ratio also depends upon the operating brightness level and, using a biasing technique on the unselected co-ordinates, allows at least 1 000 spot positions to be photographically recorded in 30 sec. The accuracy of a graphical record so produced is of a very high order, as it is dependent upon a master negative which has been mechanically fabricated. Furthermore, the long-term stability is excellent and any replacement panels will be identical.

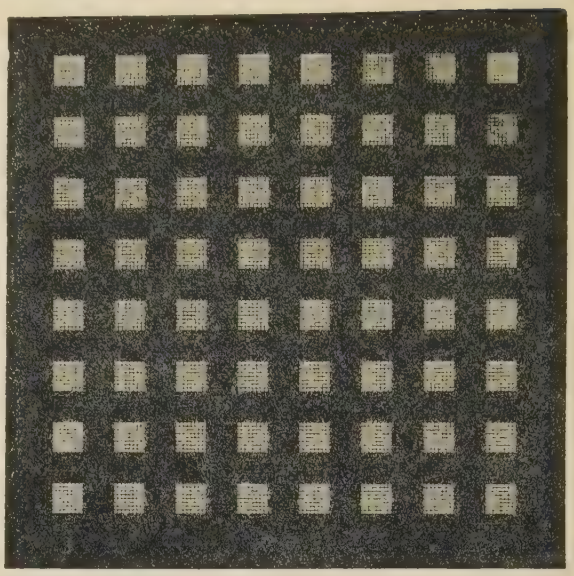


Fig. 14.—Symmetrically illuminated prototype array.

(8) ACKNOWLEDGMENTS

The authors are indebted to Professor F. C. Williams for his advice and encouragement. They also wish to thank Mr. D. C. Jeffreys, who constructed and tested the panels, and Mr. D. H. Smith, who assisted in the experimental assessment.

(9) REFERENCES

(1) 'Improvements relating to Electroluminescent Screens' (British Provisional Patent No. 12168: 1954) General Electric Company of America.

(2) DESTRIAU, G.: 'Experimental Studies on the Action of an Electric Field on Phosphorescent Sulphides', *Journal de Chimie Physique*, 1936, **33**, p. 620.
'The New Phenomenon of Electrophotoluminescence and its Possibilities for the Investigation of Crystal Lattices', *Philosophical Magazine*, 1947, **38**, pp. 700, 774 and 880.

(3) PAYNE, E. C., MAGER, E. L., and JERMOE, C. W.: 'Electroluminescence. A New Method of Producing Light', *Illuminating Engineering*, 1950, **45**, Part 2, p. 668, and *Sylvania Technologist*, January, 1951, p. 2.

(4) HENDERSON, J. J.: 'Electroluminescence', *Research*, 1955, **8**, p. 219.

(5) ENNOS, A. E.: 'Highly Conducting Gold Films prepared by Vacuum Evaporation', *British Journal of Applied Physics*, 1957, **8**, p. 113.

(6) ALFREY, G. F., and TAYLOR, J. B.: Discussion on 'The

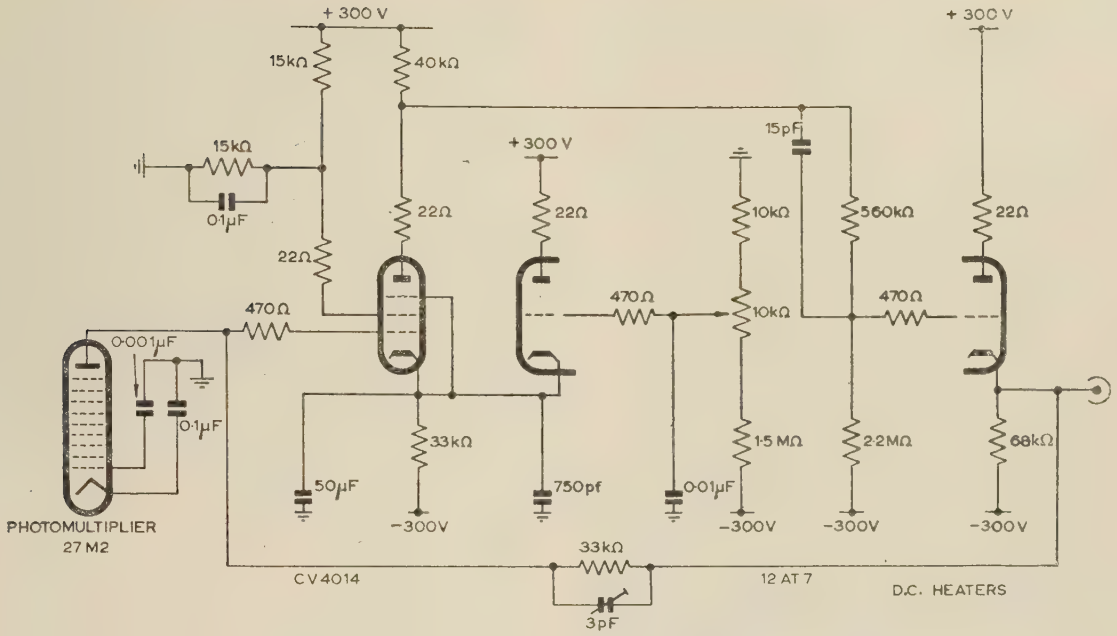


Fig. 15.—Photomultiplier current amplifier.

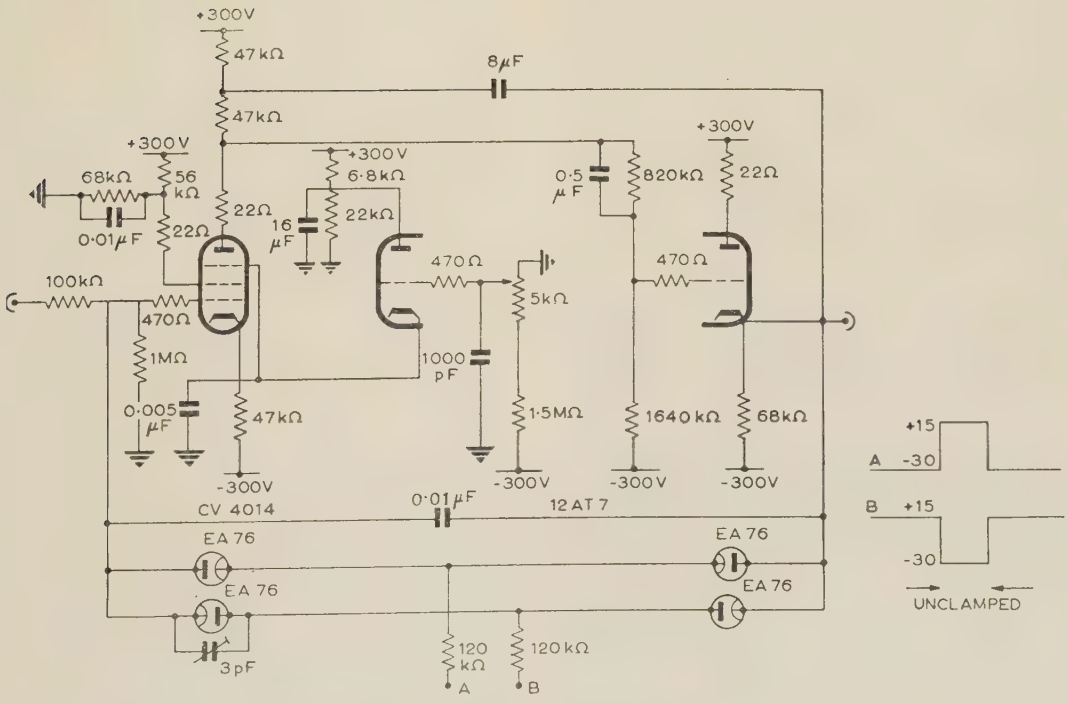


Fig. 16.—Integrator and clamp circuit.

Mechanism of Electroluminescence of Zinc Sulphide', *British Journal of Applied Physics*, 1955, **6**, Suppl. 4, p. S44.

(7) KLASENS, H. A., ZALM, P., and DIEMER, G.: 'Characteristics of Electroluminescent Cells', *Proceedings of the International Commission on Illumination, Zürich*, 1955, **1**, 2.1.1, N–K, p. 1–17 (published by Bureau Central Cie, Paris).

(8) ZALM, P., DIEMER, G., and KLASENS, H. A.: 'Electroluminescent Zinc Sulphide Phosphors', *Philips Research Reports*, 1954, **9**, p. 81.

(9) ZALM, P.: 'The Electroluminescence of ZnS Type Phosphors', *ibid.*, 1956, **11**, p. 354.

(10) DESTRIAU, G.: 'Brightness Waveforms in Electroluminescence', *British Journal of Applied Physics*, 1955, **6**, Suppl. 4, p. 49.

(11) WAYMOUTH, J. F., and BITTER, F.: 'Experiments on Electroluminescence', *Physical Review*, 1954, **95**, p. 941.

(12) HAAKE, C. H.: 'Build-up of Electroluminescent Brightness', *Journal of Applied Physics*, 1957, **28**, 2, p. 245.

(10) APPENDIX: PHOTOMULTIPLIER, AMPLIFIER, INTEGRATING AND CLAMP CIRCUITS (Figs. 15 and 16)

Circuit diagrams are shown in Figs. 15 and 16. The photomultiplier was used for its considerable amplification and subsequent circuit simplification. Care was taken so that the effect of ageing of the multiplier did not invalidate the experimental results, the overall sensitivity being checked by the standard light source technique. The multiplier anode signal is fed into an anode-follower/cathode-follower circuit to provide current amplification and low output impedance. The circuit is conventional, but it may be indicated that as the first stage accepts negative signals the anode-follower is only lightly biased. The integrator accepts positive signals and is biased accordingly; because of this the compensating triode is operated with a low anode voltage to increase the mutual conductance for a given anode current. The bootstrap connection between the cathode-follower and the anode load of the integrating stage enhances the low-frequency response. The clamp resets the input grid and output potentials, the d.c. level of the output having first been preset so that its potential is zero when the integrator input is earthed.

DISCUSSION BEFORE THE MEASUREMENT AND CONTROL SECTION, 5TH NOVEMBER, 1957

Mr. W. S. Elliott: The three papers deal with different ideas and techniques for improving the design and usefulness of electronic computers, and it is interesting to inquire whether they point a path in which technique is going. In making such an assessment, we might ask whether a technique is an improvement from the point of view of lower production cost of computers of given capacity, or of making their use easier, or of increasing their serviceability.

The object of the paper by Dr. Wilkes, Mr. Renwick and Dr. Wheeler is given in Section 1 as the simplification of overall structure of the control unit, whatever the machine order code. When the idea was first published by Dr. Wilkes in 1951, a facility to change a machine order code completely might have seemed highly desirable. Dr. Wilkes now considers that codes will not be changed after a machine is passed out.

Perhaps the authors would say whether they envisage machines with two or even more matrices, e.g. one for scientific and one for business operations, selected by switching. I would think that the enormous advantage of this design method is that it is straightforward to manufacture and to maintain. Comments by the authors on the ease or otherwise of 'knitting' the matrix would be welcome.

The authors say that most of the discussion is equally applicable to the design of the control units of machines of other types. Would they say how the method would apply to the control units of serial binary and perhaps of parallel-bit serial character machines? Do they believe that some advantages would be achieved?

Referring to the paper by Messrs. Maclean and Aspinall, it is known that Dr. Svoboda has been very interested in the possible use of the residue-number system, but there appears to be no publication by him. Could the authors tell us if he is actually using this number system?

The paper might profitably have given some consideration to the possible use of the residue-number system applied to the whole of the number representation in a parallel binary machine, e.g. in the design of a digital computer for real-time control purposes, in which numbers are not large.

So far as the general improvement in the art is concerned, it is a little difficult to assess the value of a design for an individual adder which is given out of context of a complete machine.

It is generally thought that core techniques offer permanent and trouble-free circuits. However, the techniques used by the authors involves several windings per core, each of 40 turns of 48 s.w.g. wire, through an aperture of $1\frac{1}{4}$ mm diameter. Is it significant that the authors state in the paper that 'the only components which have failed are several of the cores which developed open-circuit windings shortly after installation'?

This paper has served a very important function in drawing attention to the residue-number system. It would be interesting to hear more on the advantages of using that system in a complete computer.

Turning to the paper by Dr. Kilburn, Dr. Hoffman and Mr. Hayes, I am interested in the need for graphical output of the high accuracy afforded by the technique described. I wonder if an accuracy greater than that which can be provided by a cathode-ray tube and analogue/digital convertor can be made use of by visual assessment. The authors say that with the cathode-ray-tube method a complete trace of 1 000 spot positions can be completed in a few seconds, and I think there are equipments available which will complete a trace of that number of spots in less than one second.

Dr. D. J. Truslove: I shall confine my remarks to the paper by Dr. Kilburn, Dr. Hoffman and Mr. Hayes. The authors state that 1 000 spot positions may be recorded in 30 sec, compared with the 700 microsec required by a commercial cathode-ray-tube output recorder. The equipment has a 21 in tube for visual observation and a 7 in tube, whose accuracy is better than 0·1% full-scale, for photographic recording. This accuracy is probably sufficient for an analogue output.

A recent paper* describes a circuit which offers considerable saving in c.r.t. systems. The new equipment has maintained an accuracy of 0·1% over 39 hours.

A fast computer works in a time scale of microseconds. For one particular computer, 30 sec represents over one million additions of two 36-bit binary numbers, at a cost of about 15s. Thus it appears desirable that the electroluminescent output unit should be run as an off-line recorder, rather than directly from a computer. Have the authors considered interleaved programming, so that the machine could be fully occupied while a graphical output was being obtained?

The authors claim that the main advantage of the unit lies in its high accuracy. I would like them to explain in what circumstances this accuracy is required. There appear to be two main uses for an analogue output: first, to locate singular points in a solution, and secondly, to provide semi-technical illustrations.

* SMURA, E. J.: 'A Binary-weighted Current Decoder', *IBM Journal of Research and Development*, October, 1957.

During the Digital Computer Convention, Mr. Tizard* mentioned that the National Physical Laboratory were designing a new digital-analogue plotter. Perhaps we could have news of their progress.

In Fig. 12, curve (*b*) indicates that the integrated light output (i.l.o.) intermittent operation may exceed the i.l.o. continuous operation. The reason for this is not obvious and no explanation is given.

Mr. D. M. Taub: Referring to the paper by Dr. Wilkes, Mr. Renwick and Dr. Wheeler, in the control method described, a register is followed by a decoder which gives a signal on one out of *n* leads in the case of the diode matrix system, or switches one out of *n* cores, where *n* is the total number of order steps. The decoder is followed by two coders, a 'gates' coder, which passes the signal to the required combination of gates, and a 'next step' coder, which passes it to the combination of stages in the register corresponding to the following step.

The method is elegant and completely general, but is not the most economical where diode matrices are used. Possible economies occur mainly in the 'next step' coder. For the most part, the operation steps follow one another in a definite sequence, and the register may thus be arranged as a counting circuit. The 'next step' coder is then required only where there is a change in the normal sequence. The method is shown in Fig. A,

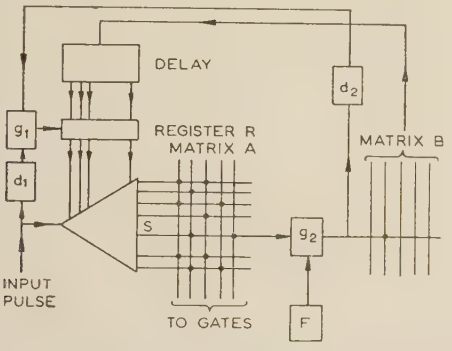


Fig. A.—A control system using a counting register and diode matrices.

which corresponds to Fig. 1 in the paper. The input pulse passes through the decoder and matrix A to operate the required combination of gates in the normal way. Where the next step is not determined by the state of the arithmetic circuits, the input pulse also passes through the delay circuit d_1 and gate g_1 to the counting input of the register. Lead S corresponds to a step where there may be a change of sequence depending on the state of a bistable circuit F in the arithmetic unit. On this step, if F contains a '0' there will be no output from g_2, and the register will therefore count as before. If F is at '1', g_2 will produce an output signal which, after passing through delay circuit d_2, inhibits g_1, so preventing the register from counting. The signal from g_2 also passes through matrix B and a delay circuit to set the register in the manner described in the paper.

Another way in which economies can often be made is to divide the gates into groups, and as far as possible arrange the order code so that a group is associated with a particular stage in the register. That stage may then be wired directly to the gates concerned without the intervening decoding and coding.

Comparing the diode and ferrite-core systems described in the paper, it can be seen that the effect of a coder diode is obtained merely by threading a wire through an already existing core, the core forming part of the decoder. The fact that this is done more often than strictly necessary is of little consequence, considering the complete generality and flexibility that result.

* 'Discussion on 'Computer Input and Output, including Analogue-Digital Conversion', *Proceedings I.E.E.*, April, 1956, 103 B, Suppl. 3, p. 448.

Mr. K. L. Smith: In the paper by Dr. Wilkes, Mr. Renwick and Dr. Wheeler, the sequential nature of the steps involved in the execution of an order suggests that simplification would result from wiring the cores 'in line' instead of in a matrix. The arrangement I have in mind is based on the magnetic shift-register. The sequence would commence by inserting a first pulse into the appropriate stage of the shift register, and the sequence of steps would be provided by this and the following stages as they were successively set and reset by the marker pulse and shifting drive. The same number of cores is required as for the matrix, but the drive is reduced to two valves; the selection is simpler, since each core feeds the next in line and is reset by one shift winding. The only limit to the number of cores driven by one drive valve is set by the delay and back-e.m.f. caused by the leakage inductance of the shift windings, but these can be reduced by increasing the drive current.

Any of the branching methods described for the matrix can be used in the shift-register.

Regarding the paper by Messrs. Maclean and Aspinall, the design of re-standardizing logical elements seems unduly influenced by the original ferrite-core devices, since most employ 2-beat operation involving distinct 'set' and 'clear' sequences and so can provide only one logical operation in the basic timing period. This normally restricts their use in binary-coded decimal and sterling adders, where the carry propagation problem causes more logical elements to be used as simple stores, to maintain correct phase relationships, than to perform the actual logic. The addition-table technique provides an elegant solution to the carry propagation problem, but what are the difficulties of extending its use to multi-radix operation? I suggest there is a need for a universal re-standardizing logical element from which the output is available almost immediately on the presentation of the input, so allowing several logical operations in the basic timing period and eliminating the phasing delays. Any logical device involving summation of flux or charge, either of which involves a subsequent resetting, is clearly not acceptable. Simple diode logic and the transistor 'nor' type circuit have suitable characteristics and could possibly be combined in a new element.

Mr. D. W. Willis: I refer to the paper by Dr. Wilkes, Mr. Renwick and Dr. Wheeler.

There is obviously much interest in the method of sequencing described, and in preference to the generalized address system, in which the address of the next wire to be stimulated can be chosen without restriction and in which the conditional jump is a simple extension, many speakers choose the particular system in which the address is normally advanced by one and the conditional jump is a special case.

It would be interesting to know in what proportion of the steps in the system described in the paper the next address is fixed, and in what proportion the next address is conditional.

Dr. T. B. Tomlinson: I should like to comment on the paper by Dr. Kilburn, Dr. Hoffman and Mr. Hayes with regard to the brightness discrimination, in which I also have been interested.

The authors improve the discrimination by operating at low voltage and reduced brightness. The scheme which I have suggested includes an additional layer of granular semi-conducting material, e.g. cadmium sulphide activated with chlorine, between the electroluminescent layer and the top (non-transparent) electrodes. This additional layer is rather thicker than the electroluminescent layer and acts substantially as a non-linear resistance, the current through it increasing approximately as the fourth power of the voltage across it. The electroluminescent layer acts almost as a perfect capacitor, independent of applied voltage.

It is arranged that the resistance of the semi-conducting layer is at least equal to, and preferably greater than, the reactance of the electroluminescent layer when the full operating voltage is

applied. Where the 'one-third' voltage appears across unwanted intersections the resistance of the intermediate layer is much greater because of the fourth-power law, and the voltage appearing across the electroluminescent layer is greatly reduced. The new voltage ratio tends to a limiting value of 3^4 as the thickness of the intermediate layer is increased; a much improved brightness ratio results.

The authors suggest that there might be some possibility of using evaporated layers in the future. May I, without wishing to be discouraging, give a note of warning? Early experiments show that evaporated phosphor layers are only weakly fluorescent unless they are given a subsequent heat treatment to improve the crystal structure. This necessitates a substrate with a thermal coefficient of expansion matching that of the fluorescent layer and may add considerable difficulty to electrode construction. Also, the phenomenon of electroluminescence takes a different form in single-crystal material and does not depend on thickness in the same way as for powder layers.

THE AUTHORS' REPLIES TO THE ABOVE DISCUSSION

Dr. Wilkes, Mr. Renwick, and Dr. Wheeler (*in reply*): We discuss in the paper a number of schemes for the design of the control unit of an otherwise conventional digital computer. These are to be regarded as alternative to the more usual schemes. We have never ourselves attached great importance to the idea of providing separate control matrices for use in solving different classes of problem. Mr. Elliott mentions the possibility of providing one control matrix for scientific applications and one for business applications. We feel, however, that the requirements for the two sets of applications are not fundamentally different, and that a complete set of instructions covering both can be provided with a single control matrix of sufficient size.

The operation of 'knitting' even a large matrix is not a great task compared with the total amount of work involved in the construction of a digital computer. The making of minor corrections and alterations has turned out to be surprisingly straightforward, even in a matrix containing 32 rows and 32 columns.

The ideas described in the paper can be applied to the design of a control system for a serial machine. In this case, the outputs of matrix A are not used to control directly gates in the machine. Instead they control the passage of waveforms synchronized to the clock-pulse generator and having appropriate starting times and durations.

Various modifications of the systems we have described are possible. There are, in addition, a variety of other systems, some of which are mentioned in the discussion, which also make use of diodes or magnetic cores, although most of these do not perhaps go as far as ours in meeting the objectives specified in the third paragraph of the introduction to our paper. One should beware of attempting to simplify the control matrix of the system shown in Fig. 1 (for example) at the expense of providing equivalent diode switching circuits elsewhere.

Several speakers have assumed that during the execution of an instruction the steps follow one another in an invariable order, except for the rare occurrence of a conditional or branching point. This would be so if conditional steps occurred only in such operations as multiplication, where the action depends on whether a certain digit in a number is a 0 or a 1. In practice, however, conditional steps are also introduced in order to enable sub-sequences of operations which occur in more than one instruction to be performed by the same set of cores in the control matrix. For example, in the control matrix of Edsac 2 the same division sub-sequence is used both for fixed-point division and for floating-point division, switching to and from the subsequence being effected by means of conditional cores. The result is that, in the control matrix of Edsac 2, two out of every seven intersections are conditional, which means that one out of every two cores is conditional. We doubt, therefore, if the scheme suggested by Mr. Taub would lead to any great saving, especially in view of the fact that a counter must be provided to control the decoding tree instead of a simple register.

Messrs. Maclean and Aspinall (*in reply*): We understand that Dr. Svoboda is considering the logic of a computing machine based upon the residue-number system, but do not know whether he is contemplating the construction of such a machine.

The dilemma associated with the application of this number system to computing machines is as follows: if one is to represent large numbers the number of moduli chosen must also be large, and the difficulty of identifying these numbers increases as the number of moduli increases. In any computer based on this system there must be a limit to this number of moduli which is used, and special provision will have to be made to identify numbers which are greater than the lowest common multiple of the moduli.

Whilst the residue-number system has certain advantages in this exercise, it is felt that the system is inflexible and not immediately applicable to the design of a general computing machine.

The technique of placing the windings on the core by hand requires a certain amount of experience. Early cores were prone to failure because of faulty winding, most of the failures occurring immediately after installation. When the adder was completed it underwent tests and was in operation for a considerable time. No failures due to faulty 48 s.w.g. wires occurred during this time.

Dr. Kilburn, Dr. Hoffman and Mr. Hayes (*in reply*): In reply to Mr. Elliot and Dr. Truslove we should like to distinguish between three states, namely

(i) Maximum speed of an output equipment in isolation.
(ii) Maximum speed of output obtained when the same output equipment is connected to a computer.
(iii) Speed of output really required.

In state (i) we realize that spot positions can be recorded at the rate of about one per microsecond on a cathode-ray tube. We should, however, be very surprised if this were maintained in a commercial equipment in state (ii), and even more surprised if this rate were really required. If, as we believe, it is not required at present, Dr. Truslove's costing at 15s. per photograph and his remarks about off-line recording and interleaved programming are irrelevant.

It is perhaps worth noting that the speed of 1 000 spot positions in 30 sec, quoted for the present electroluminescent panel, is equivalent to 200 decimal digits per second.

Both Mr. Elliot and Dr. Truslove are critical of the accuracy feature of the electroluminescent panel. Yet it would appear from their remarks that both are concerned about accuracy, Mr. Elliot being content with 'that which can be provided by a cathode-ray tube and analogue/digital convertor' and Dr. Truslove with 'an accuracy of 0·1% over 39 hours'. We, in our turn, are content from the engineering and maintenance point of view to forget accuracy now that the master negative has been made, and from the users' point of view are content that a given point is always in the same place.

In reply to Dr. Tomlinson, we find ourselves in agreement with his interesting suggestion. Recently considerable progress has been made on evaporated phosphor layers, so we do not despair too soon of the difficulties which we may encounter.

621.372.832.43 : 621.372.821

The Institution of Electrical Engineers
Paper No. 2402 R
Aug. 1957

BROAD-BAND SLOT-COUPLED MICROSTRIP DIRECTIONAL COUPLERS

By J. M. C. DUKES, M.A., Associate Member.

(The paper was first received 8th March, and in revised form 12th June, 1957. It was published in August, 1957, and was read before the Radio and Telecommunication Section 13th November, 1957.)

SUMMARY

The paper describes a new design technique for directional couplers in a printed strip-above-ground microwave transmission system (microstrip). The two transmission lines are mounted back to back and coupled through slots in the common ground plane. The coupling factors of isolated slots of different shape are calculated and are found to be in good agreement with measured results. A specific design of 3 dB coupler is described in which the coupling mechanism consists of 50 transverse slots graded in length. By grading the phase velocity in one of the two lines a power split equal within 1 dB over the band 2 800–4 300 Mc/s is achieved.

(1) INTRODUCTION

The adoption of printed strip transmission lines,[1,2] as a replacement for conventional hollow waveguide or coaxial cable, has necessitated the development of a variety of components, including directional couplers.[3] So far, only two types have been described in the literature. In the first,[4] shown in Fig. 1(*a*), the two strip conductors are run side by side in close proximity over a distance of several centimetres, the necessary coupling occurring through fringe-field interaction. This arrangement has the merit of extreme simplicity but at the same time suffers from several disadvantages. In the first place, it is difficult to obtain tight coupling unless the conductors are very close (e.g. less than 0·5 mm apart) or the coupling section very long (e.g. several wavelengths). Secondly, there is evidence that the coupled section can resonate in a balanced twin-strip mode. This produces erractic variations in directivity similar to the long-slot effect in long-slot directional-couplers in waveguide. A method for overcoming this has been suggested but there still remains the third and most pertinent objection, namely the difficulty of controlling both the phase-change coefficient and the coupling factor in a prescribed manner.

The second type of coupler, illustrated in Fig. 1(*b*), is that known as the branched-arm coupler.[2] This has certain advantages over the parallel-line coupler. However, if a high value of directivity over a large bandwidth is required, the number of branches becomes rather large (e.g. more than 3) and the impedance of the outer arms tends to become higher than can conveniently be realized in this or any similar form of construction (e.g. more than 150 ohms).

These considerations suggested that the arrangement shown in Fig. 1(*c*), where the two lines are mounted back to back, might have many advantages. Coupling is achieved by means of slots printed in the common ground plane. By a suitable choice as to size and position of the coupling holes, a wide variation in coupling strength is possible. Moreover, according to the shape of the hole, either forward or reverse directivity can be obtained.

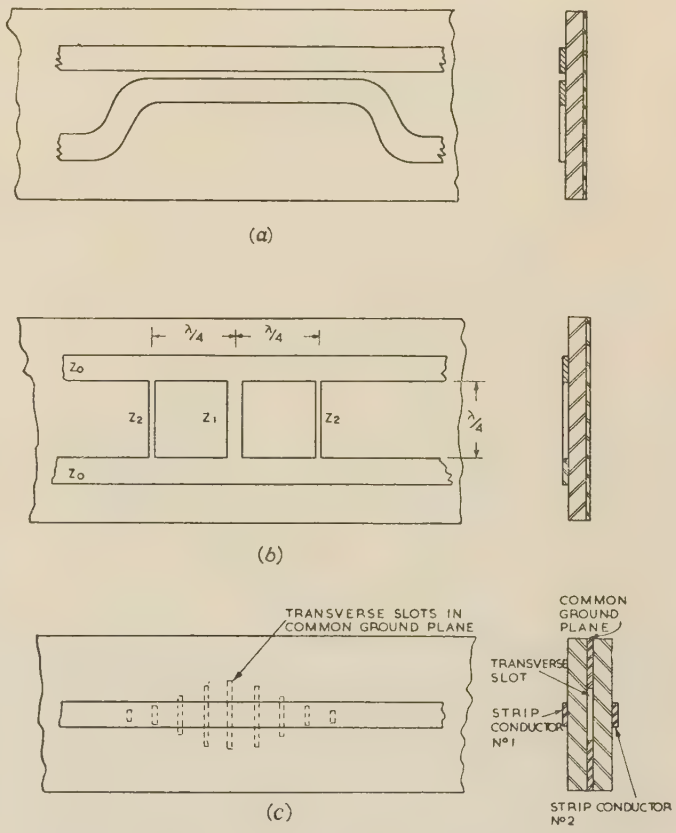


Fig. 1.—Different forms of microstrip directional coupler.

(*a*) Parallel-line coupler.
(*b*) Branched-arm coupler.
(*c*) Back-to-back slot-coupled coupler.

(2) SINGLE-SLOT COUPLING

(2.1) Outline of Theory

The procedure adopted for calculating the coupling coefficient of a single slot is described in greater detail in Section 7. The approach is quite classical[5] and differs from the technique used for hollow waveguide in one major respect only. Since no solution in analytic form has as yet been derived for the electromagnetic field in a practical dielectric-sheet-supported microstrip line, it is necessary to resort to one or other of several possible artifices. For a centrally located slot whose transverse dimension is small compared to the width of the strip conductor, the method first proposed by A. A. Oliner[7] has been adopted. This is illustrated in Fig. 2; the transmission line of Fig. 2(*a*) is replaced by a fictitious line, as shown in Fig. 2(*b*), in which the dimension h is maintained but the width b' is chosen to make the characteristic impedance of the two lines identical. The impedance of the true line may be found in Reference 6, and the impedance of the fictitious line is given quite simply by

$$Z_0 = \frac{h}{b'}(\mu_d/\epsilon_d)^{1/2} \quad . \quad . \quad . \quad . \quad . \quad (1)$$

Mr. Dukes is with Standard Telecommunication Laboratories Limited.

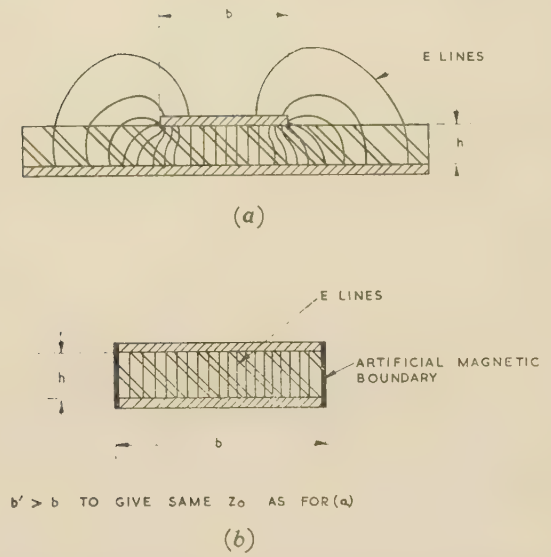


Fig. 2.—Approximation involved in calculating coupling factor.
(*a*) Actual microstrip line.
(*b*) Artificial line.
$b' > b$, to give same Z_0 as for (*a*).

where μ_d and ϵ_d are the magnetic permeability and dielectric constant of the dielectric medium. The general theory of energy radiation through small holes[5] may now be applied in the usual manner. Provided that certain limitations are observed in regard to the thickness of the common ground plane and the size of the holes, the degree of approximation involved in this method of analysis is small.

For slots whose transverse dimension is equal to or larger than the strip conductor width it is necessary to take into account the transverse variation in field across the slot. In the absence of any more precise data it has been found convenient to employ the numerical data given in Fig. 12 of Reference 6, which were obtained from measurements made on an electrolytic tank. In using these results three approximations were involved. First, the data used relate to a line immersed in a homogeneous dielectric medium and therefore neglect the modifying effect of the dielectric interface. Secondly, there exists a small discrepancy in regard to the actual proportions of the line. Thirdly, in order that the classical small-hole theory may be applied in a straightforward manner, it is necessary to assume that the field across the slot is constant with an effective value equal to the average value. Fortunately it is rarely necessary in practice to use slots having a transverse dimension greater than the strip conductor width by more than about 50%, in which case the approximations involved are not in themselves likely to cause any very serious discrepancy.

(2.2) Experimental Results

The following experimental technique was used. Two specimens of microstrip line were prepared from copper-clad p.t.f.e.-impregnated Fibreglass, each involving a straight run of conductor approximately 14 cm long. A certain area of copper foil, larger than the largest slot, was stripped off each ground plane. The two lines were then clamped firmly back to back, with a thin piece of copper foil in between, in which had been punched a slot of the desired size and shape. Locating holes and screws were provided to ensure that the component pieces were always replaced in the same relative positions each time the dimensions of the coupling slot were modified.

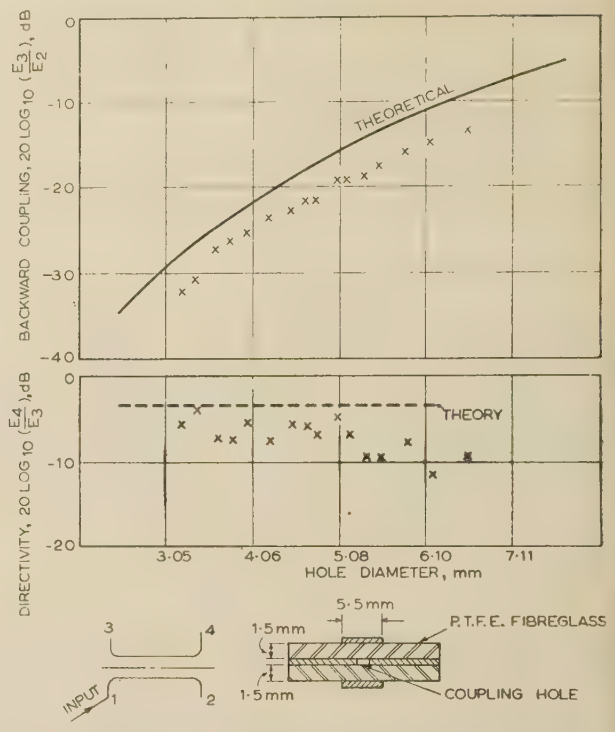


Fig. 3.—Coupling through a single symmetrically located circular slot.
To minimize effect of dissipation coupling is taken as E_3/E_2 rather than E_3/E_1. Test frequency, 4 000 Mc/s.

Fig. 3 shows the results for a single symmetrically located circular slot measured at 4 000 Mc/s. The stray coupling was measured by reducing the slot size to zero, under which condition the coupling was too small to be measured, i.e. it was at least 60 dB down. Also shown in Fig. 3 is the theoretically computed curve calculated from formulae derived in Section 7, but in principle applicable only to very small holes. The agreement is reasonable in view of the limitations in both the experimental technique and the theoretical analysis.

An experiment was also conducted in which the angle between the two microstrip lines was varied between 0° and 90°, thus

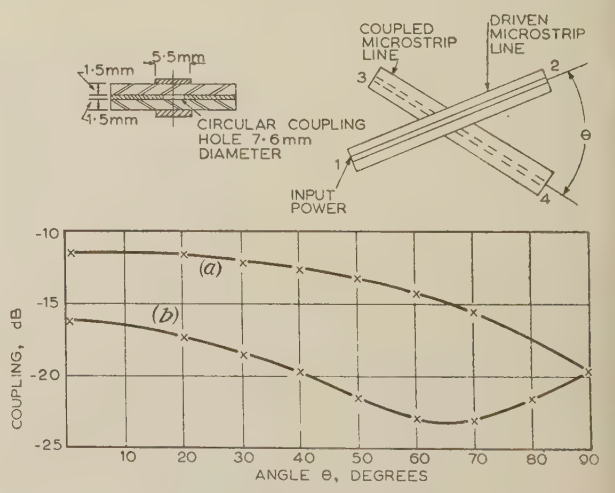


Fig. 4.—Bethe–Hole coupler.
To minimize effect of dissipation, coupling is referred to E_2 rather than E_1. Test frequency, 4 000 Mc/s.
(*a*) Backward coupling, 20 $\log_{10}(E_3/E_2)$.
(*b*) Forward coupling, 20 $\log_{10}(E_4/E_2)$.

constituting a Bethe–Hole coupler.[8] The results are illustrated in Fig. 4. Contrary to theory, no very sharp minimum was observed, but this may be because of the excessive size of the coupling hole necessary for a satisfactory measurement of directivity. Moreover, minimum transmission does not occur at quite the angle predicted from eqn. (16) (Section 7), namely 79°.

Experimental results for transverse rectangular slots are given in Fig. 5. These have the advantage over circular slots that, for

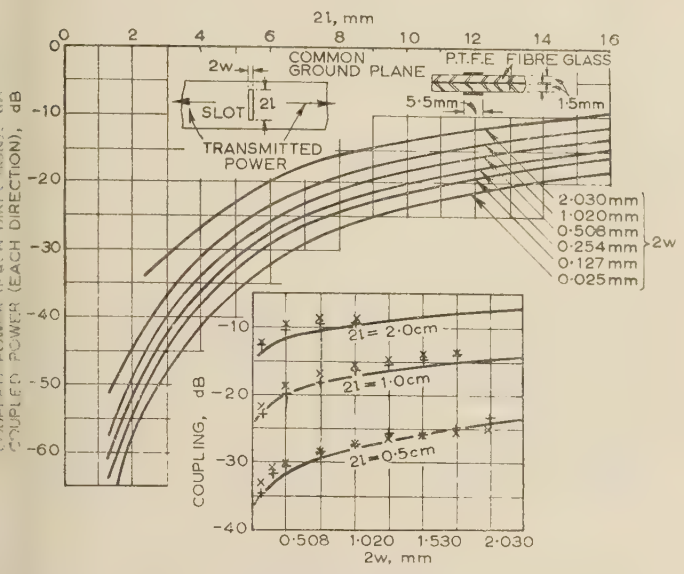


Fig. 5.—Coupling through a transverse rectangular slot.

× Forward wave
\+ Reverse wave
} measured points.

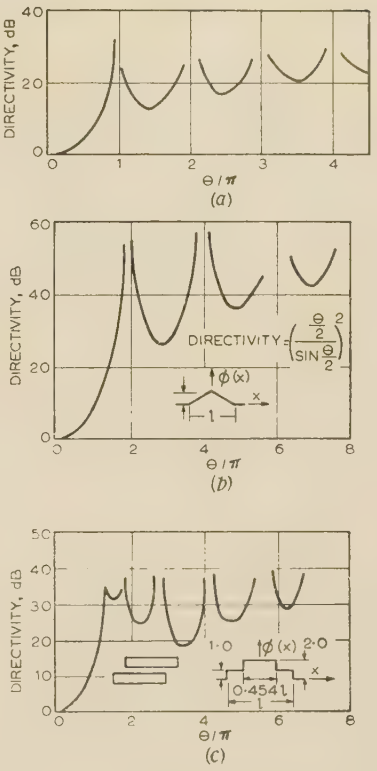


Fig. 6.—Some typical directivity functions.

(*a*) Uniform coupling.
(*b*) Linear taper coupling.
(*c*) Two superimposed uniform couplings.

practical purposes, coupling takes place as a result of excitation by the transverse component of the magnetic field only; excitation by the transverse component of the electric field or by any spurious longitudinal components is equally quite negligible. As a result the directivity is small, and generally in the forward direction, which is an advantage in so far as multi-slot couplers are concerned. Resonances may occur if the transverse dimension of the slot is too large, but in general such wide slots are not required. Taking into account certain approximations in the theoretical analysis and limitations in the measuring technique, the agreement indicated in Fig. 5 is again quite reasonable.

(3) MULTI-SLOT DIRECTIONAL COUPLERS

(3.1) Outline of Theory

As is well known, high values of both directivity and coupling may be obtained by the use of a multiplicity of slots. S. E. Miller has treated this problem from a very general point of view,[10] and it is his approach that is used in this paper (Section 8). First, a continuous coupling mechanism described by some relatively simple analytic function is assumed. Miller has tabulated curves of directivity for a number of such functions, three of which are shown in Fig. 6. The parameter θ is the mean electrical length of the coupled section, in the present instance (TEM propagation) a linear function of frequency. Assuming that moderate directivities between 20 and 30 dB are considered satisfactory and that the operating bandwidth is limited to a ratio of approximately 2 : 1, the third distribution has the advantage of the shortest coupling length. The next step is to replace the continuous coupling function by a series of discrete elements. Two methods are possible, namely equal-strength couplings with variable spacing, or variable-strength couplings with constant spacing. The first case, treated in some detail by Miller,[10] is of special interest when the slots have to be prepared by a machining process, in that only one type of tool is required. However, with printed-circuit techniques no such limitation exists, and it is preferable to take advantage of the higher efficiency offered by the second alternative. From data given by Miller it appears that quantization into about eight discrete couplings will give a reasonable approximation to the desired directivity curve, provided that the electrical length at the highest frequency does not exceed 4π. It is preferable, however, to use a rather larger number; this is of little consequence since the number of printed slots has virtually no effect on the final cost of the component.

The required coupling strength of the individual slots is calculated from eqn. (25) of Section 8.2. For reasons outlined in Section 2.2 a transverse rectangular slot is preferred. Control of coupling may be achieved by varying either the longitudinal or the transverse dimension of the slot, but the latter is simpler.

When calculating the desired size of the slot it is essential to take into account the variation of coupling with frequency. Eqns. (11) and (12) of Section 7 show that the coupling factor of a single slot is a simple linear function of frequency. From eqn. (23) of Section 8.2 it follows that in this case the overall coupling factor will vary sinusoidally with frequency. For a small coupling coefficient this variation is 6 dB per octave, which is rather more than can be tolerated in many applications.

Fortunately there exists a relatively simple method of reducing this frequency variation. It is shown in Section 8.2 that if the two transmission lines are made to have different phase velocities the energy exchange is still periodic, but the period is modified and a complete exchange of energy never occurs. For the case

of uniform coupling and uniform phase velocities the principal parameter is

$$q = \tfrac{1}{2}|\beta_1 - \beta_2|/c \quad . \quad . \quad . \quad . \quad . \quad (2)$$

where β_1 and β_2 are the phase-change coefficients of the two lines (radians per metre) and c is the distributed coupling coefficient in the same units.

Neglecting dissipation, optimum performance will be obtained when q is chosen to ensure the following condition at mid-band:

$$cl\sqrt{(1 + q^2)} = \pi/2 \quad . \quad . \quad . \quad . \quad (3)$$

where l is the length of the coupled section. This has been plotted in Fig. 7. Application of this method to the design of a

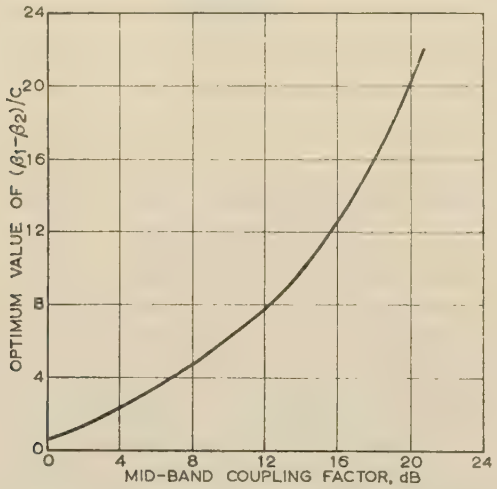


Fig. 7.—Design curve for optimum phase compensation.

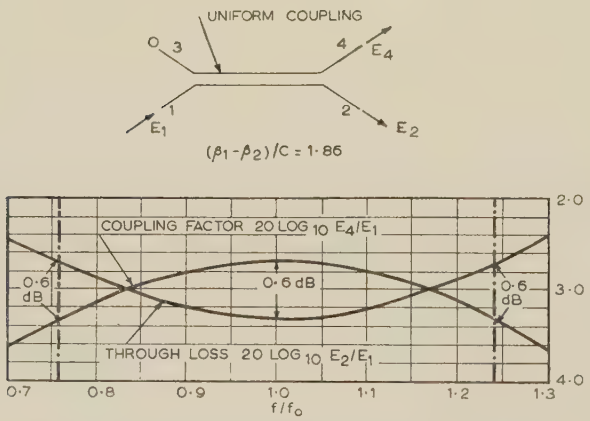


Fig. 8.—Optimum phase-compensated 3 dB coupler.
Theoretical performance for a given bandwidth ratio of 1·68 : 1.

3 dB directional coupler with a bandwidth ratio of 1·64 : 1 is shown in Fig. 8, from which it can be seen that the total variation can be reduced, in theory at least, from −2·0 + 1·3 dB to only 0·3 dB.

Uniform coupling is of little practical interest, since it leads to rather low values of directivity unless the coupled section is exceptionally long. However Miller's solution for uniform coupling is also valid for the non-uniform case, provided that the factor q is maintained constant along the coupled section. The coupling characteristic as a function of frequency will then be identical to that for a uniform section having the same values of $\int_0^l q dx$ and $\int_0^l c dx$. The change in $|\beta_1 - \beta_2|$ may be abrupt or gradual, but some care is necessary to avoid provoking undesirable reflections since the directivity of the coupler will be seriously impaired thereby.

(3.2) Experimental Results

In constructing these experimental couplers it was found convenient to use a master slotted section comprising 100 equally spaced transverse rectangular slots over a length of 20 cm, the slot dimensions being 3 cm in the transverse direction and 0·254 mm longitudinally. The master pattern was obtained by photographing an accurately machined ebonite block. The ground plane of one line is etched with this pattern and then a copper foil mask is used to restrict the coupling so as to obtain the desired coupling function. For reliable results it is necessary to solder the mask in position, carefully removing all surplus solder. The ground plane is stripped off the second line and the two microstrip lines are clamped firmly back to back. Other experimental techniques have been tried but this proved the quickest and most reliable. Once the design has been finalized, however, it is preferable to etch the exact pattern of slots required and thus obviate the use of a mask.

Attention will be concentrated on the 3 dB coupler illustrated n Fig. 9. It consists of two microstrip lines mounted back to

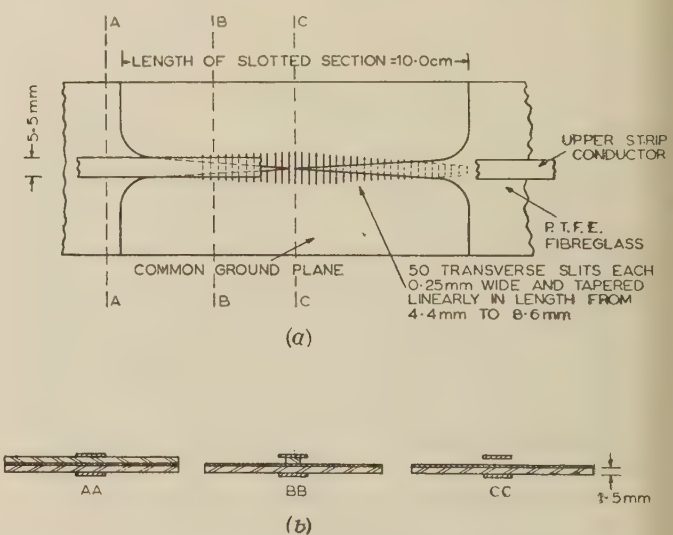


Fig. 9.—Experimental phase-compensated 3 dB coupler.
(*a*) Plan view of upper surface.
(*b*) Successive cross-sections.

back, one of which has its dielectric support progressively cut away so that at the centre of the section the phase-change coefficient is virtually that for an air-spaced line. The coupling slots are likewise tapered in magnitude so that the factor q is maintained roughly constant along the length of the section. The overall length of the coupled section is only 10 cm.

The design parameters, based on a mid-band frequency of 3 500 Mc/s, are as follows. The optimum value of $q = 0{\cdot}93$, determined from Fig. 8, gives a mid-band value of integrated coupling strength $C_l = \int_0^l c(x)dx = 1{\cdot}15$ radians. For the fully supported Fibreglass line, $\beta_1 = 110$ rad/m, and at the centre of the other line $\beta_2 = 73$ rad/m, which is the free-space value.

Hence, at the centre of the coupled section, $|\beta_1 - \beta_2| = 37$ rad/m. Assuming a linear taper for both the coupling factor and the difference of phase-change coefficients, the necessary length of the coupled section works out at 11·6 cm, with a peak coupling of 20 rad/m at the centre. From Fig. 6(*b*), this gives an expected directivity of 26 dB at mid-band, with higher values over the rest of the band.

In practice it was found convenient to replace the linear tapers of the electrical parameters by linear physical tapers. The slot width plotted as a function of distance along the coupled section was found to be very nearly linear, except at the extremities where the width decreases very rapidly. The simplest approximation to this function is therefore a raised linear taper. In the case of the velocity taper little data existed and the profile was selected on the basis of a qualitative knowledge of the field.

The initial choice for the transverse dimension of the slot was 7·6 mm at the centre tapering to 2·6 mm at the ends of the section. This arrangement gave too little coupling, particularly at the low end of the band, and some empirical adjustments were therefore necessary. The final slot lengths were 6·6 mm at the centre and 4·4 mm at the end. These adjustments are quite simple. From a study of the phase relationships in the coupled section it is clear that a *pro rata* increase in coupling over the entire section gives a more or less proportionate increase in coupling at all frequencies. On the other hand, increasing the coupling at the ends of the section increases the coupling at the low-frequency end of the spectrum, and if carried far enough will eventually cause a reduction in coupling at the upper end of the spectrum.

In view of the approximations involved both at the design stage and during the subsequent empirical adjustments it is not possible to make an exact comparison with theory. However, the results of Fig. 10 are sufficiently close to the predicted values

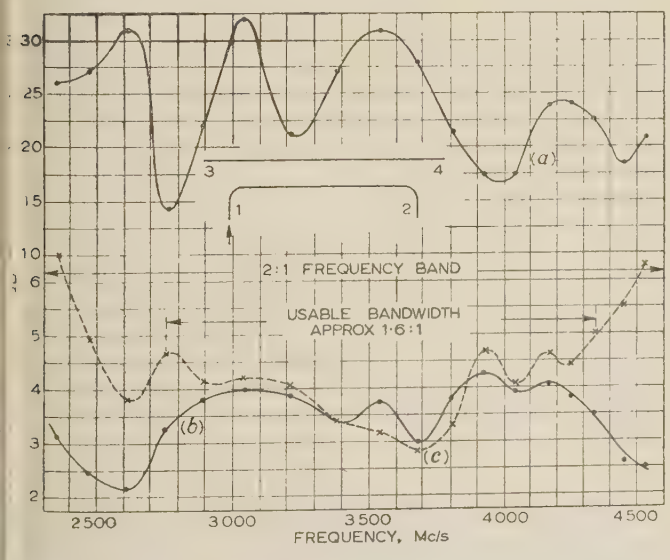


Fig. 10.—Experimental phase-compensated 3 dB coupler.
Measured performance over 2 : 1 band.

to justify the general technique. The rather large dips in directivity are perhaps a little disappointing, but could no doubt be overcome by the use of a longer coupling section.

(4) CONCLUSIONS

The theoretical and experimental results described provide a sound basis for the design of satisfactory directional-couplers in microstrip. Full use is made of the non-dispersive properties of microstrip to ensure satisfactory performance over a very wide band of frequencies. In addition, the coupler has advantages over conventional waveguide in regard to size and ease of fabrication. No attempt has been made to ensure other than a very nominal value of directivity, e.g. 20 dB, but there is no doubt that this could be improved upon quite substantially. As an experimental exercise the 3 dB coupler has provided an interesting confirmation of the theoretical principles developed by Miller,[10] Cook,[11] and Louisell,[12] which are applicable also to other types of transmission medium. The method of calculating the radiation through slots in the ground plane should prove useful in the design of microstrip antenna systems.

(5) ACKNOWLEDGMENTS

The author is indebted to several of his colleagues, in particular Mr. D. A. Williams, who constructed the apparatus and made most of the measurements. The paper is published by permission of Standard Telecommunication Laboratories Limited.

(6) BIBLIOGRAPHY

(1) Grieg, D. D., and Engelmann, H. F.: 'Microstrip—A New Transmission Technique for the Kilomegacycle Range', *Proceedings of the Institute of Radio Engineers*, 1952, **40**, No. 12, p. 1644.

(2) Fromm, W. E.: 'Characteristics and Some Applications of Stripline Components', *Transactions of the Institute of Radio Engineers*, March, 1955, Vol. MTT-3, No. 2.

(3) Montgomery, C. G., Dicke, R. H., and Purcell, E. M.: 'Principles of Microwave Circuits', M.I.T. Radiation Laboratory Series No. 8 (McGraw-Hill, New York, 1948).

(4) Arditi, M.: 'Characteristics and Application of Microstrip for Microwave Wiring', *Transactions of the Institute of Radio Engineers*, March, 1955, Vol. MTT-3, No. 2.

(5) Surdin, M.: 'Directional Couplers in Waveguide', *Journal I.E.E.*, 1946, **93**, Part III A, p. 725.

(6) Dukes, J. M. C.: 'An Investigation into some Fundamental Properties of Strip Transmission Lines with the Aid of an Electrolytic Tank', *Proceedings I.E.E.*, Paper No. 1991 R, May, 1956 (**103** B, p. 319).

(7) Oliner, A. A.: 'Equivalent Circuits for Discontinuities in Balanced Strip Transmission Lines', *Transactions of the Institute of Radio Engineers*, March, 1955, Vol. MTT-3, No. 2, p. 134.

(8) Montgomery, C. G.: 'Technique of Microwave Measurements', M.I.T. Radiation Laboratory Series No. 11 (McGraw-Hill, New York, 1948).

(9) Dukes, J. M. C.: 'The Application of Printed Circuit Techniques to the Design of Microwave Components' (see page 155).

(10) Miller, S. E.: 'Coupled Wave Theory and Waveguide Applications', *Bell System Technical Journal*, 1954, **33**, No. 3, p. 661.

(11) Cook, J. S.: 'Tapered Velocity Couplers', *ibid.*, 1955, **34**, No. 4, p. 807.

(12) Louisell, W. H.: 'Analysis of the Single Tapered Mode Coupler', *ibid.*, 1955, **34**, No. 4, p. 853.

(7) APPENDIX: RADIATION THROUGH SMALL HOLES

The general theory of radiation through small slots is well established, hence no attempt will be made to derive the basic expressions used in the following analysis which may be found elsewhere.[3, 5] In the present instance the analysis can be limited to the situation illustrated in Fig. 11, where two transmission lines 1 and 2 intersect at an angle $\theta + \phi$, the common surface

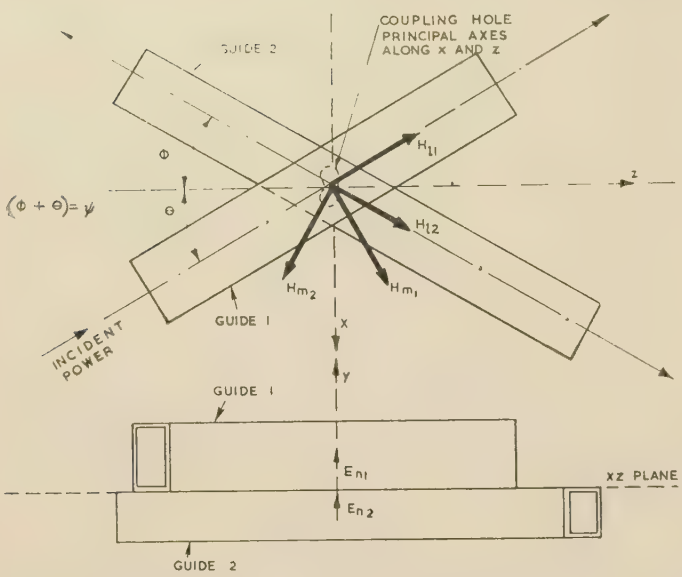


Fig. 11.—Slot coupling of two TE guides.
(*a*) Decoupling, 20 $\log_{10} |E_3/E_1|$.
(*b*) Through loss, 20 $\log_{10} |E_2/E_1|$.
(*c*) Coupling, 20 $\log |E_4/E_1|$.

of intersection lying in the *zx*-plane. For simplicity the principal axes of the slot are assumed to coincide with the *x*- and *z*-axes. Moreover, it will be assumed that each guide can support only one mode.

The symbols employed have the following meanings:

M_x = Magnetic polarizability of the slot along the *x*-axis.
M_z = Magnetic polarizability of the slot along the *z*-axis.
P = Electric polarizability of the slot (perpendicular to the plane of the hole).
H_{m1}, H_{m2} = Transverse components of the magnetic field in each guide at the point $x = y = z = 0$.
H_{l1}, H_{l2} = Longitudinal components of the magnetic field in each guide at the point $x = y = z = 0$.
E_{n1}, E_{n2} = Electric field in each guide at the point $x = y = z = 0$ assumed perpendicular to the guide wall.

The components H_{m1}, H_{l1}, E_{n1} are the unperturbed values which would exist in the absence of the slot. The components H_{m2}, H_{l2}, E_{n2} must likewise satisfy the normal mode equations for guide 2.

It can be shown that the amplitude transmission coefficients A and B of the forward and reverse waves in the second guide are given by

$$A = \frac{j\pi}{\lambda S}(m_m H_{m2} + m_l H_{l2} - P E_{n1} E_{n2}) \quad . \quad . \quad . \quad (4)$$

$$B = \frac{j\pi}{\lambda S}(-m_m H_{m2} + m_l H_{l2} - P E_{n1} E_{n2}) \quad . \quad . \quad (5)$$

where

$$m_m = m_x \cos\phi - m_z \sin\phi \quad . \quad . \quad . \quad . \quad (6)$$

$$m_l = m_z \cos\phi + m_x \sin\phi \quad . \quad . \quad . \quad . \quad (7)$$

and

$$m_x = M_x(H_{m1} \cos\theta - H_{l1} \sin\theta) \quad . \quad . \quad . \quad (8)$$

$$m_z = M_z(H_{l1} \cos\theta + H_{m1} \sin\theta) \quad . \quad . \quad . \quad (9)$$

The term S is a normalizing factor proportional to the Poynting vector, and is given by

$$S = \int \boldsymbol{E} \times \boldsymbol{H} \boldsymbol{n} d\sigma \quad . \quad . \quad . \quad . \quad . \quad (10)$$

the integral being taken over the cross-section of the second guide, where $\boldsymbol{n}$ is a unit vector in the *y*-direction.

For two TEM transmission lines the H_{l1} and H_{l2} components are zero. Making the further simplification that $\theta = \phi = \psi/2$, we have

$$A = \frac{j\pi}{\lambda S}[H_{m1} H_{m2}(M_x \cos^2\psi/2 - M_z \sin^2\psi/2) - P E_{n1} E_{n2}] \quad . \quad (11)$$

$$B = \frac{j\pi}{\lambda S}[-H_{m1} H_{m2}(M_x \cos^2\psi/2 - M_z \sin^2\psi/2) - P E_{n1} E_{n2}]. \quad (12)$$

Values of M_x, M_z and P have been worked out for several types of slot.[3,5] For example, for a circular slot,

$$M_x = M_z = \tfrac{4}{3} r^3, \; P = \tfrac{2}{3} r^3 \quad . \quad . \quad . \quad . \quad (13)$$

where r is the radius of the slot. Another case of importance is the long thin slot lying along the *x*-axis. This may be considered as an elliptical slot of large eccentricity with major semi-axis l and minor semi-axis w, where $2l$ and $2w$ are the length and width of the true slot. The solution for this is

$$M_x = \frac{2}{3} \frac{l^3}{\log(4l/w) - 1} \quad . \quad . \quad . \quad . \quad (14)$$

$$M_z = P = \frac{2}{3} l w^2 \quad . \quad . \quad . \quad . \quad . \quad (15)$$

In the case of TEM transmission lines where the field solution is given in parametric form, application of eqns. (3) and (4) is relatively straightforward. However, no analytic solution as yet exists for the practical microstrip line of Fig. 2(*a*). For this reason it is necessary to resort to the artifice proposed by A. A. Oliner[7] and described in Section 2.1, and eqn. (1).

For the commonly employed dimensions of $b = 5{\cdot}5$ mm and $h = 1{\cdot}5$ mm, and p.t.f.e. Fibreglass as the dielectric, the value of Z_0 may be taken as approximately 47 ohms. In this case b' works out at 7·4 mm, assuming a dielectric constant of 2·65. For a circular hole and a frequency of 4000 Mc/s, eqns. (11) and (12) reduce to

$$A = 4{\cdot}09 j r^3(2 \cos\psi - 0{\cdot}377) \quad . \quad . \quad . \quad (16a)$$

$$B = -4{\cdot}09 j r^3(2 \cos\psi + 0{\cdot}377) \quad . \quad . \quad (16b)$$

where r is in centimetres. Application of these equations is illustrated in Fig. 3.

For a transverse slot of small dimension, i.e. $2l < b$, eqns. (14) and (15) may be substituted directly in eqns. (11) and (12). Moreover, provided that $2l \gg 2w$, and ψ is small, M_z and P may be neglected in comparison with M_x. Hence, for $\psi = 0$, we have

$$A = B = 6{\cdot}44 \frac{l^3}{\log(4l/w) - 1} \quad . \quad . \quad . \quad (17)$$

where l and w are in centimetres.

When the length $2l$ of the slot becomes comparable with or greater than the breadth b of the strip conductor, the above formula is no longer accurate. In order to calculate the coupling due to larger slots use was made of data previously obtained on an electrolytic tank, as mentioned in Section 2.1. Since the magnetic field is no longer constant along the slot, it is necessary to make an assumption as to the effective value. In the absence of any better assumption a simple average was taken. The value of S was then taken to be that of the fictitious line as previously described. The results of this calculation are shown in Fig. 5.

(8) APPENDIX: COUPLED WAVE THEORY

(8.1) Loose Coupling Theory

Consider two lossless transmission lines coupled by some unspecified form of coupling mechanism. Let l be the length of the coupled section, and $k(x)$ the coupling function, defined as the ratio of the series e.m.f. per unit length induced at the point x in the coupled line by the voltage on the driven line at the same point. It will be assumed for simplicity that the coupling mechanism is non-directive, and only the principal mode can propagate in each line.

Provided that the coupling is sufficiently loose, the directivity, defined as the ratio of the forward to the backward current in the coupled line, is given by

$$\text{Directivity} = \left[\int_{-\frac{1}{2}}^{+\frac{1}{2}} k(x)dx\right]\left[\int_{-\frac{1}{2}}^{+\frac{1}{2}} k(x)\varepsilon^{j2\theta x/l}dx\right]^{-1} \quad . \quad (18)$$

where $$\theta = -\pi l(1/\lambda_1 + 1/\lambda_2) = -\tfrac{1}{2}l(\beta_1 + \beta_2) \quad . \quad . \quad (19)$$

and λ and β are, respectively, the wavelengths and phase-change coefficients as denoted by the subscripts.

A number of these discrimination functions have been tabulated,[10] three of which are illustrated in Fig. 6. When the continuous coupling function is simulated by a series of equally spaced point couplings, the directivity function becomes periodic in θ/π. It is symmetrical about $\theta/\pi = \frac{1}{2}(n-1)$ with the first zero located at $\theta/\pi = n - 1$. The number of point couplings required to give a reasonable approximation to the continuous function therefore depends on the operating bandwidth and the variation in coupling over this band.

(8.2) Tight Coupling

Consider now the case where a significant amount of power is abstracted from the driven line by the coupled line. In order to simplify the problem the following assumptions will be made. First, let the characteristic impedance of each line be normalized to unity, the power carried in the line is then given by $|E|^2$, where E is the amplitude of the forward wave in that line. Secondly, let the coupling per wavelength be relatively small. In this case it is necessary to consider only waves travelling in the forward direction, and hence only the first derivative appears in eqns. (20) and (21) below. Thirdly, let it be assumed that the coupling mechanism is reciprocal. It can be shown that the space variation of the wave amplitude is given quite closely by

$$\frac{dE_1}{dx} = -(\gamma_1 + k)E_1 + kE_2 \quad . \quad . \quad . \quad (20)$$

$$\frac{dE_2}{dx} = kE_1 - (\gamma_2 + k)E_2 \quad . \quad . \quad . \quad . \quad (21)$$

where γ_1 and γ_2 are the uncoupled propagation coefficients of the two lines, and k is the coupling coefficient. Assuming the coupling mechanism to be free of dissipation, k must be imaginary. In this case it is convenient to introduce a real quantity c defined by $k = jc$.

The first and simplest solution occurs when the two transmission lines are identical, i.e. $\gamma_1 = \gamma_2 = \gamma$. We then have, for the condition $E_1 = 1$ and $E_2 = 0$ at the input to the coupled section,

$$E_1 = \cos C_l\varepsilon^{-j(C_l+\gamma l)} \quad . \quad . \quad . \quad . \quad (22)$$

$$E_2 = j\sin C_l\varepsilon^{-j(C_l+\gamma l)} \quad . \quad . \quad . \quad . \quad (23)$$

where C_l is the integrated coupling strength for a section of length l and is given by

$$C_l = \int_0^l c(x)dx \quad . \quad . \quad . \quad . \quad . \quad . \quad (24)$$

It will be noted that this solution does not depend on the manner in which the coupling function $c(x)$ varies along the coupled section. The other features of this solution are that the power tends to vary cyclically from one line to the other, and at $C_l = \pi/2$ all the power in line 1 is transferred to line 2. Moreover, the waves in the two lines are 90° out of phase.

Where multiple discrete couplings are used the following formula applies for C_l:

$$C_l = \sum_r \text{arc sin } c_r \quad . \quad . \quad . \quad . \quad . \quad (25)$$

where c_r is the magnitude of the rth coupling. The order in which the couplings occur along the line is of no significance except in so far as the directivity is concerned.

In practice, $c(x)$ is a function of frequency also. For the slot couplings described in Section 7 the coupling is, in theory at least, a simple linear function of frequency. From eqns. (22) and (23) it is clear, therefore, that the minimum rate of change of coupled energy in line 2 will occur when $C_l = \pi/2$, which is the condition corresponding to complete power transfer to line 2. This is evidently a condition of little practical interest. Fortunately, as has been shown by Miller, it is possible to overcome this limitation by employing two transmission lines having different phase velocities.

To cover this case it is useful to define a new parameter, q, by

$$q = \tfrac{1}{2}(\beta_1 - \beta_2)/c \quad . \quad . \quad . \quad . \quad . \quad (26)$$

Subject to the restriction that q be constant along the coupled section, and with the simplification that the attenuation is the same for each line, the solution eqns. (20) and (21) is

$$E_1 = \bar{E}_1\varepsilon^{-[\alpha l + jC_l + j\frac{1}{2}(\beta_1-\beta_2)l]} \quad . \quad . \quad . \quad . \quad (27)$$

$$E_2 = \bar{E}_2\varepsilon^{-[\alpha l + jC_l + j\frac{1}{2}(\beta_1-\beta_2)l]} \quad . \quad . \quad . \quad . \quad (28)$$

where

$$\bar{E}_1 = \cos[(1+q^2)^{1/2}C_l] - jq(1+q^2)^{-1/2}\sin[(1+q^2)^{1/2}C_l] \quad . \quad . \quad . \quad . \quad . \quad (29)$$

and $$\bar{E}_2 = (1+q^2)^{-1/2}\sin[(1+q^2)^{1/2}C_l] \quad . \quad . \quad (30)$$

From this it is apparent that the exchange of energy is still periodic, but with a modified period. Furthermore, a complete exchange of energy never occurs, the maximum amplitude of the wave in line 2 being given by the coefficient $(1+q^2)^{-1/2}$.

For TEM transmission lines β is a simple linear function of frequency. Assuming that c varies linearly with frequency also, q must be constant with respect to frequency. From eqn. (30) it is clear that under these circumstances the amplitude of the wave in line 2 varies sinusoidally with frequency. The rate of change is a minimum at $(1+q^2)^{1/2}C_l = \pi/2$, and the optimum performance will therefore be achieved if the coupled section is so designed that this condition occurs at the mid-band frequency (see Fig. 7). The choice of $c(x)$ will be determined by the directivity function desired, and the functions $\beta_1(x)$ and $\beta_2(x)$ must then be chosen to make $q(x)$ a constant.

In the most general case $q(x)$ may also vary along the coupled section, and it is to be expected that by a judicious choice of this function the frequency sensitivity of the coupler may be even further reduced. This more general case, however, is difficult to deal with, since eqns. (20) and (21) combine to give a

linear differential equation of the second order with non-constant coefficients:

$$\frac{d^2E}{dx^2} + P(x)\frac{dE}{dx} + Q(x)E = 0 \quad . \quad . \quad . \quad . \quad (31)$$

where $P(x)$ and $Q(x)$ are functions of $\beta_1(x)$, $\beta_2(x)$, $c(x)$ and their differentials. As is well known, the solution to this equation cannot be presented in closed form except for specified functions for $P(x)$ and $Q(x)$.

Cook[11] and Louisell[12] have considered this case and hav shown that one solution exists giving constant coupling over a indefinitely wide bandwidth. This solution has the disadvantag that the length of the coupled section must be many wavelength From a consideration of what happens when several uniform sections having different values of q are connected in cascade it is clear that the coupling may be made very nearly constan over bandwidths of the order of 2 : 1 without the need fo excessively long coupled sections.

[The discussion on the above paper will be found on p. 180.]

621.372.821.049.75

The Institution of Electrical Engineers
Paper No. 2401 R
Aug. 1957

THE APPLICATION OF PRINTED-CIRCUIT TECHNIQUES TO THE DESIGN OF MICROWAVE COMPONENTS

By J. M. C. DUKES, M.A., Associate Member.

(*The paper was first received* 8*th March, and in revised form* 12*th June*, 1957. *It was published in August*, 1957, *and was read before the* RADIO AND TELECOMMUNICATION SECTION 13*th November*, 1957.)

SUMMARY

A brief résumé is given of the basic theory of strip transmission lines including unwanted effects such as spurious mode transmission and radiation. A variety of materials and different printed-circuit techniques have been tried and the relative merits of each are discussed. A technique of measurement is described and details are given of two pieces of specially developed measuring equipment in microstrip, namely an adjustable short-circuit and a wide-band precision attenuator. A number of new broad-band components covering the range 2 500–4 200 Mc/s are described; they include a reverse-phase hybrid-ring circuit, a back-to-back slot-coupled 3 dB directional coupler, and a new type of low-pass filter. The paper concludes with a description of an r.f. head for two-way communication equipment used in propagation testing.

(1) INTRODUCTION

The paper has been written with two aims in view. The first is to provide information concerning a range of new microwave components, details of which have not previously been published. These components have the common feature that the transmission lines employed are in planar form and are fabricated by means of a printed-circuit technique. Neither the form nor the method of fabrication is new, however,[1–4] and the novelty of these components lies rather in the manner in which the technique has been exploited and extended.

In developing these components free use has been made of each of the several well-known types of strip transmission line.[1, 5, 6] As a result a fairly broad view of the strip-line technique has been obtained. The second aim of the paper, therefore, is to present an assessment of the relative merits of the different systems in any intended application. It is felt that this will be of particular value to the less specialized reader.

It should be emphasized that printed-circuit techniques can also be used as an ancillary technique in the fabrication of conventional waveguide components. For example, where one wall of the waveguide carries a complex arrangement of slots, the wall can, under certain circumstances, be replaced by an etched circuit. For various reasons, however, this technique has found little application so far.

(2) ELECTRICAL PROPERTIES OF STRIP TRANSMISSION LINES

(2.1) Types of Strip Transmission Line

The three conductor configurations in current use are illustrated in Fig. 1. They are equivalent, respectively, to open-wire balanced pair, wire above ground, and coaxial cable. The practical versions of these three systems are shown in Fig. 2. The first will be referred to as 'balanced strip line', and the second is now universally known as 'microstrip'.[1–3] The nomenclature for the lines of Fig. 2(*c*) and (*d*) is somewhat confused; Fig. 2(*c*) has been referred to variously as 'flat-strip',[4] 'sandwich', 'triplate'[6] and 'shielded parallel-plate line'.[7] Of these alternatives, 'triplate' would appear the most elegant, and will be used throughout this paper. Fig. 2(*d*) is a modification whereby

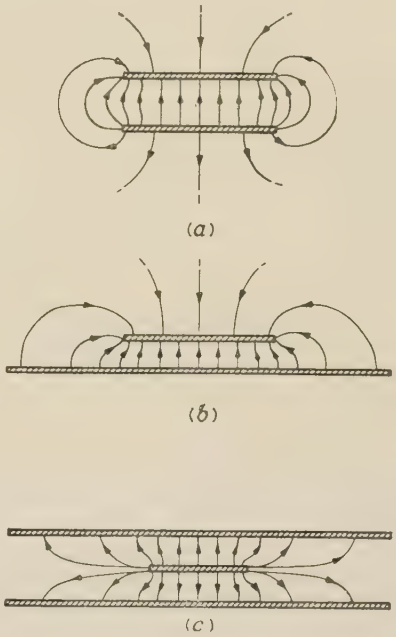


Fig. 1.—Geometries of the three principal strip transmission lines.
(*a*) Balanced line. (*b*) Strip above ground. (*c*) Triplate.

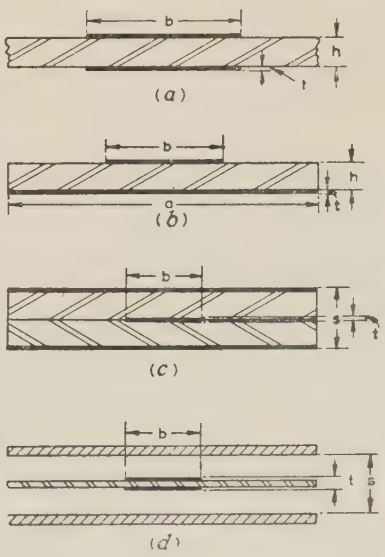


Fig. 2.—The four practical types of strip transmission line.
(*a*) Balanced line. (*b*) Microstrip. (*c*) Triplate. (*d*) High-*Q* triplate.

Mr. Dukes is with Standard Telecommunication Laboratories Limited.

the dielectric is confined to the low-field region, thus allowing higher Q-factors than can be obtained with the three other forms. One of the strips can be omitted, but this disturbs the symmetry of the cross-section and as a result a small increase in dielectric loss occurs.[8] In this paper the line of Fig. 2(*d*) will be designated 'high-Q triplate line', although some writers refer to it as 'high-Q strip line'.[5] The latter designation is not very satisfactory, since the term 'strip line' has acquired a more general connotation covering the entire range of strip transmission lines.

The respective merits of these principal forms are summarized briefly in Table 1, whilst Sections 2.2 and 2.3 discuss some of in Reference 9. The accuracy deteriorates progressively in the directions of increasing ϵ_d/ϵ_0 and decreasing b/h, but in the central region (e.g. around $b/h = 2{\cdot}0$ and $\epsilon_d/\epsilon_0 = 3$) measured results agree with the computed figures within about 2%.

A somewhat similar method could no doubt be employed to calculate the impedance and phase velocity of high-Q triplate line, Fig. 2(*d*), although so far no attempt has been made to do so. Referring to Fig. 1(*c*), it will be seen that the dielectric sheet lies along the plane of symmetry where the electric gradient is zero, in theory at least. Hence it is to be expected that for thin dielectric sheets the perturbation will be considerably less than

Table 1

COMPARISON OF DIFFERENT STRIP-LINE SYSTEMS

	Balanced strip line	Microstrip	Triplate	High-Q triplate
Ease of fabrication	2	1	3	4
Bulk and weight	1	1	2	3
Ease of testing	2	1	4	3
Ease of analysis	3	3	1	2
Dissipative losses	2	2	3	1
Radiation losses	2	2	1	1
Topological limitations ..	Difficult to feed from unbalanced transmission line.	Phase reversals not possible, but ground plane easily tapered away to give balanced line.	E-plane series junctions difficult. Phase reversals impossible. Discontinuities must in principle be symmetrical.	
Appropriate fields of application.	Broad-band hybrid junctions involving phase reversals and/or series junctions.	Broad-band low-Q components.	Broad-band components and medium-Q filters (e.g. high-pass and low-pass).	Broad- and narrow-band components. Filters generally including high-Q resonators.
Use to date*	Minor.	Extensive.	Considerable.	Extensive.

Numbers represent order of merit.
* This assessment refers to the situation within the strip-line field. It is not possible, as yet, for a variety of reasons, to predict how far strip line will replace existing coaxial and waveguide technique.

these points in greater detail. It should be noted that the lines shown in Figs. 1 and 2 are not wholly inclusive but they are believed to represent those forms which are most likely to find any extensive application.

(2.2) Theory of Operation and Principal Electrical Properties

The simplified theory of strip transmission lines has been discussed in some detail in a previous article.[9] Propagation in the transverse electromagnetic (TEM) mode is assumed, and the analysis therefore reduces to a solution of the 2-dimensional electrostatic field in a simple cross-section. This assumption is entirely valid for the air-spaced lines of Fig. 1, and also for the low-Q triplate line of Fig. 2(*c*), provided that in the latter case the dimension a is considerably larger than the strip conductor width b. However, with Figs. 2(*a*), (*b*) and (*d*) the wave experiences a perturbation due to presence of the longitudinal dielectric interface. As a result the wave acquires longitudinal components of both the electric and magnetic vectors and the dominant mode therefore ceases to be pure TEM. In practice, however, provided that the dimensions b and h remain a small fraction of a wavelength (measured in the dielectric), the dispersive effect is very small and any TEM solution which takes into account the modified capacitance will give values of phase velocity and effective impedance which are in close agreement with measured results.

This latter technique has been used quite effectively in the case of the two transmission lines shown in Figs. 2(*a*) and (*b*). The generalized curves for impedance and wavelength, illustrated in Figs. 3 and 4, were constructed using the method described

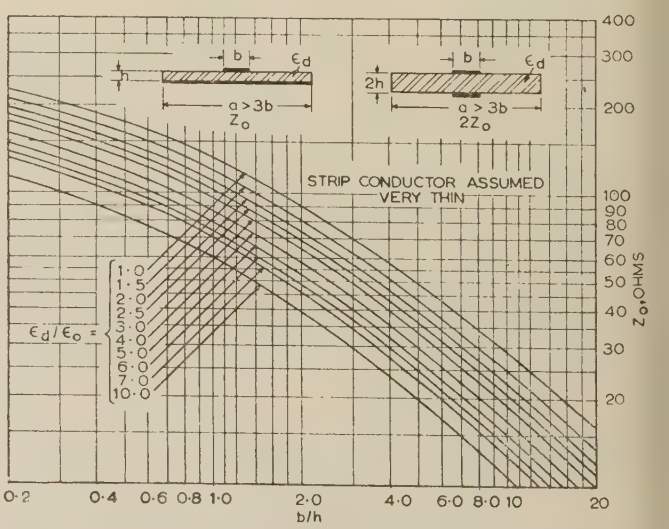


Fig. 3.—Characteristic impedance of microstrip and balanced line.
Pseudo-TEM solution.
Strip conductor assumed to be very thin.

with microstrip, and this is borne out in practice. For a t/b ratio of the order of 0·2 or less the effect on the characteristic impedance as calculated by Cohen,[7] Fig. 5, is generally quite negligible.[8] The effect on the phase velocity is more marked[8] and must be taken into account in the design of filters and similar components. For the materials and dimensions commonly

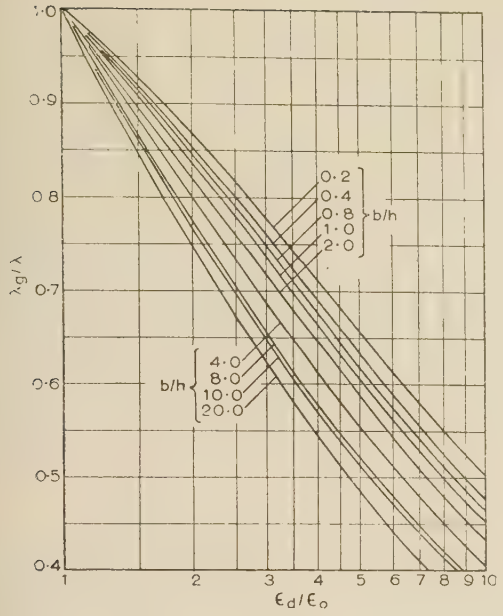


Fig. 4.—Wavelength in microstrip and balanced line.
Pseudo-TEM solution.

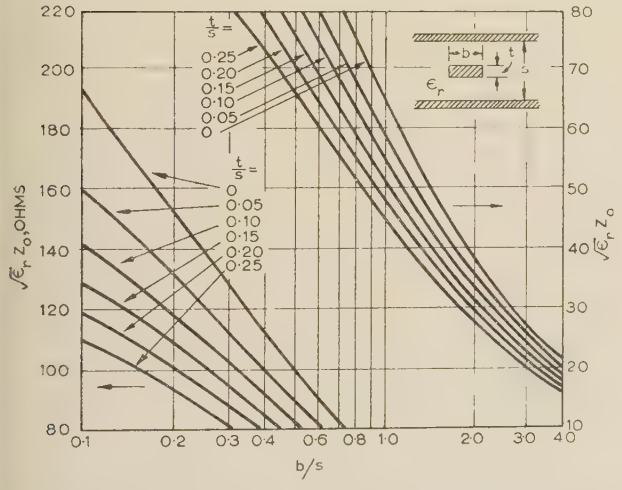


Fig. 5.—Characteristic impedance of triplate line.
Calculated by S. B. Cohn and reproduced from *Transactions of the Institute of Radio Engineers*, 1955, MTT-3, p. 119.

employed the reduction in phase velocity ranges between 5 and 15%.

The conductor losses in strip transmission lines have also been investigated theoretically.[9] For the dimensions generally employed the loss does not differ very greatly from that of the simple parallel-plate line having infinitely wide plates, namely

$$\alpha = R(Zh)^{-1} \text{ nepers} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (1)$$

where R = R.F. skin resistance of conductor, ohms per unit length.
Z = Wave impedance in dielectric.
$= 377\,\epsilon_0/\epsilon_d'$ ohms.
h = Ground plane spacing, fractions of unit length.
ϵ_d' = The effective dielectric constant.

For Figs. 2(*a*) and (*b*) it is best to take $\epsilon_d' = \epsilon_d$, where ϵ_d is the dielectric constant of the supporting sheet, whilst for high-Q triplate line, Fig. 2(*d*), a value of $\epsilon_d' = \epsilon_0$ may safely be assumed in most cases. More precise calculations[9] take into account the concentration of current at the edges of the strip conductor; the ratio of the actual loss to that given by eqn. (1) rises almost linearly with increasing h/b (i.e. as the strip narrows) and at $h/b = 1$ the loss may be as much as twice that calculated by the above simple formula. The measured attenuation[11] for a microstrip line having the dimensions $b = 5{\cdot}5$ mm, $h = 1{\cdot}6$ mm, and with polytetrafluorethylene (p.t.f.e.) Fibreglass dielectric, varies from about $\frac{1}{2}$dB/m at 1 000 Mc/s to about 3 dB/m at 10 000 Mc/s. The corresponding figures for a practical high-Q triplate line[5] will be at least one order of magnitude lower.

(2.3) Higher Modes, Spurious Modes and Anomalous Effects

In all cases it is desirable to restrict the strip conductor width to somewhat less than $\frac{1}{2}\lambda$ in the dielectric medium, otherwise the next higher mode having zero electric field at the centre of the strip conductor may propagate. The cut-off of wavelength for this mode has been determined precisely for triplate line[12] but not as yet for microstrip. Likewise it is also desirable to restrict the spacing between strip conductor and adjacent ground plane to less than $\frac{1}{4}\lambda$ in dielectric. Both these requirements are usually satisfied quite easily up to frequencies of the order of 12 000 Mc/s.

A more troublesome feature with strip transmission lines is the occurrence of spurious propagation in modes unrelated to the dominant mode, and even more serious is the possibility, under certain unfavourable circumstances, of direct radiation. In the case of triplate line these spurious modes are usually restricted by the boundary conditions to an unbalanced TEM wave between the two ground planes, or, if the sides are closed, waves of the type H_{0n}. They are stimulated either by a slight longitudinal tilt of the strip conductor or by an asymmetrical discontinuity, and will provoke radiation at the sides of the line unless the circuit is completely enclosed. Trouble due to these modes can usually be eliminated by short-circuiting the two ground planes together in the region of the discontinuity with an array of suitably disposed metal pins.[5, 6, 13]

In the case of microstrip the situation is somewhat different. For the air-spaced line with $h < \frac{1}{2}\lambda$ and $b < \frac{1}{2}\lambda$ the boundary conditions do not permit any solution other than the dominant TEM mode. A similar conclusion would also appear valid in the case of Fig. 2(*b*), provided that the width of the dielectric sheet does not exceed approximately $\frac{1}{2}\lambda$ in the dielectric, although in this case the wave must in fact possess small components of both E and H in the direction of propagation. When, as frequently happens in practice, the sheet is several wavelengths wide the situation is rather obscure. For example, it may be that a whole series of modes can propagate having properties almost identical with the dominant mode but differing field distributions in the more remote regions on either side of the strip conductor. Secondly, there is the possibility of spurious propagation by modes which can be sustained independently of the dominant mode. Assuming a flat sheet of p.t.f.e. Fibreglass $\frac{1}{16}$ in thick with foil on one side, the only known possibility at normal centimetric frequencies is the lowest-mode transverse-magnetic surface wave.[14] Discontinuities at the edge of the sheet and elsewhere will certainly cause this mode to radiate rather severely. Certain types of discontinuity in the strip conductor system, such as a shunt inductive post, will evidently radiate directly without the intermediary of a confined mode. In this connection it should be stressed that considerable care is required when assessing whether or not a certain microstrip component is radiating. Since microstrip is an open structure the unperturbed field extends (in theory) to infinity, and power

will be coupled into any circuit placed in the neighbourhood of the transmission line. Since the bulk of the energy propagates in the region between the strip conductor and the ground plane, the power coupled into a probe situated above the dielectric sheet is normally very small. An unusually large pick-up would suggest direct radiation, and this may be cross-checked by examining how the coupled power varies with the distance from the line. Decay inversely proportional to the square of the distance would suggest radiation.[6]

Despite this very unprepossessing list of difficulties the microstrip system can in practice be made to function very satisfactorily, but the field of application is usually limited to broadband low-Q components such as those described later in the paper. Within this field it evidently has many advantages over triplate lines, in regard to both experimental adjustment and subsequent manufacture.

In conclusion there is one phenomenon common to all strip transmission lines which should be mentioned. The transition from waveguide or coaxial inevitably involves a rather severe change in the field structure, and the field in the strip line near the transition therefore consists of the dominant-mode wave together with enough higher-mode waves to satisfy the boundary conditions at the transition. (Under unfavourable circumstances some radiation may also be involved.) The intensities of the higher-mode waves decrease rapidly as they are normally beyond cut-off, and at a sufficient distance from the transition only the dominant wave is left. This effect is particularly marked in microstrip, where it has been observed that the attenuation constant is a function of distance from the transition.[11] It rises rather rapidly over the first few wavelengths and eventually reaches an asymptotic limit at a distance of some 20 or more wavelengths from the transition.

(2.4) Upper Frequency Limit

For reasons already explained in Section 2.3 it is essential to keep the dimensions b and h within certain limiting fractions of a wavelength. Hence the maximum permissible physical size in cross-section will vary as f^{-1}. Substituting in eqn. (1) and recognizing that $R \propto f^{1/2}$, it will be seen that the minimum possible conductor loss per unit length varies as $f^{3/2}$. In addition, it is necessary to allow for a dielectric loss which will vary linearly with f, and for the reduction in size of the circuit in the direction of propagation which will be at the rate of f^{-1}. These considerations give a law of the form

$$\text{Minimum circuit loss} = m + nf^{1/2}\,\text{decibels} \quad . \quad . \quad . \quad (2)$$

where m and n are constants determined by the materials used and the type of strip transmission line. For high-Q triplate line m is very small, provided that a good-quality board is used, and the second term therefore dominates. With microstrip, however, the dielectric losses are more important, and the second term becomes significant only at the upper end of the centimetric band.

So far, strip-line techniques have in general been confined to the centimetric region of 1 000–12 000 Mc/s. The above considerations suggest that their use could be extended to considerably higher frequencies, e.g. Q-band, before the losses became exorbitant. In practice, the frequency limit would be set by the difficulty of manufacturing very small printed-circuit panels with a sufficient degree of precision, a difficulty, however, which is not unique to strip line.

(3) PREPARATION OF THE CIRCUIT

(3.1) Review of Printed-Circuit Techniques

Printed-circuit techniques may be classified roughly in two categories.[15–17] In the method most widely exploited to date the dielectric board—usually called the laminate—is supplied already clad with copper foil, and the process consists in removing the foil not required. The foil to be retained is protected by a resist which may be deposited in a variety of different ways, and the unwanted foil is then removed by spraying with a suitable etching solution. The method is relatively simple and when properly conducted can provide a very high standard of dimensional stability.

The second category covers methods whereby the metal circuit is deposited directly on a plain board. Amongst the better-known methods are silk-screen printing with silver ink, silver evaporation under vacuum, and electrolytic deposition on a printed graphite trace. These methods have found very limited application at microwave frequencies, principally because the conductivity and line definition are inclined to be rather poor. Moreover, the composition of silver ink tends to vary in an irregular manner, and it has been reported that this can cause a microstrip line to radiate significantly.[18]

(3.2) The Etched-Foil Technique

The first step is to prepare a master drawing of the circuit, usually several times full scale. This may be done either with pen and ink on good-quality Bristol board, or with suitable tools on scraper board, which is a white clay board coated with a thin black surface. The latter method is usually more accurate, since the scraping tool can be machined to the precise width required, and often much quicker. Where extreme accuracy is required the master can be machined from a solid block of some material such as ebonite. The master drawing or pattern is then photographed and where necessary reduced to the required size.

In order to print the resist on to the copper-clad board any one of several techniques may be used. For this class of work the most satisfactory method is the so-called photo-mechanical technique, wherein the board is coated with photo-sensitive emulsion and exposed by contact printing in a vacuum frame. The undeveloped emulsion is then washed off, leaving a protective coating of resist where the copper is to be retained.

Whichever process of depositing the resist is used the etching technique is basically the same. The most common procedure is to spray the board evenly with ferric perchloride until the desired amount of foil has been removed, and then wash and dry. Where required the etched board is finally given a flash of silver or other appropriate metal.

For experimental work or small-scale production the photo-mechanical method has many advantages. The dimensional repeatability is excellent, e.g. of the order of a fraction of 0·001 in, particularly if glass transparencies are used in place of film. On the other hand, some difficulty has been experienced due to undercutting of the strip conductor by the etchant, but provided that the process is carefully controlled the error caused by the undercut is small, e.g. less than 0·001 in for 0·001 4 in foil; it is a consistent error and can be allowed for at the design stage.

(4) MATERIALS

(4.1) General Requirements

There are certain requirements regarding materials which are common to all forms of printed-circuit applications. These include reasonable mechanical rigidity, good adhesion of metallic conductors, ease with which soldered connections can be made, ability to withstand the expected climatic conditions, and resistance to vibration and shock. These requirements are basically similar whatever the operating frequency of a printed circuit, but in the microwave case they tend to be tighter, since a small deterioration in physical condition may have a relatively severe effect on electrical performance.

In addition, there are a number of requirements which are usually of little concern at low frequencies. The dielectric board must have a very low electrical loss, as also any adhesive used to bond the conductor to the board. Values of tan δ up to 0·01 are sometimes acceptable, but values lower than 0·001 are usually preferred. The board must be of uniform thickness and composition, otherwise impedance irregularities will occur causing serious mismatch and reflection, and the dimensions and composition must remain stable under changing climatic conditions. The board should be free from internal strains, since these will release during etching and cause local warping. (For the high-Q triplate line the tolerance on warp is particularly tight.) The water absorption must be exceptionally low, otherwise the loss angle and dielectric constant will vary in service. Furthermore, an absorbent board will experience contamination during the etching process. Variations in the exact value of dielectric constant can be allowed for at the design stage, after which it must be held within tight limits. These limits will depend on the particular component, its application and the type of strip line used. Under certain circumstances a tolerance of $\pm 1\%$ might be required, but usually the designer can arrange matters so as to conform to a somewhat more realistic figure. In this respect the high-Q triplate line has a very definite advantage, since the dielectric sheet is in the weak-field region.

(4.2) Typical Materials

It is convenient to classify the dielectric materials according to whether or not the board is reinforced with Fibreglass. The unsupported board will usually consist of a high-grade material such as polyethylene or polystyrene. It has been found possible, by a controlled application of heat and pressure, to bond copper foil to these two materials without the aid of any adhesive. The bond strength is sufficient for many applications although not of the first order. The electrical properties are excellent and better than those of any materials so far tested. Moreover the copper foil presents a smooth surface which is easily etched. On the other hand, both materials have rather severe mechanical limitations: the softening temperature is rather low; polyethylene is too flexible for most microwave applications, whilst polystyrene is undesirably brittle. Neither material withstands the etching process very well since the release of internal strains may cause the board to warp. With sheets of polystyrene having a thickness equal to or less than $\frac{1}{32}$ in this has resulted in the specimen becoming crazed with fine hair cracks, to the extent that it may shatter on subsequent handling.

The latter problem can be entirely overcome by moulding Fibreglass cloth into the sheet. In order to keep the losses low it is desirable to use as little Fibreglass as possible; in practice two cloths, one each side, are sufficient. In one variation of this technique the Fibreglass cloth is printed with silver ink before moulding into the styrene board. As can be seen from Table 2, this provides a reasonably low-loss circuit with the advantages of a flush surface and very high bond strength. It will also be observed that a worthwhile improvement in r.f. losses is gained by substituting fine copper gauze for the ground plane instead of inked Fibreglass. This method of producing the circuit has not found much application, however, because of the relatively poor definition. Instead it is found preferable to mould copper foil on to the surface and print the circuit by the etched-foil technique in the usual manner. Expanded ethylene and styrene have also been tried, but it is difficult to obtain sufficient expansion in such thin sections; the result is that the losses per wavelength are of the same order as for the solid sheet, with the disadvantage of poor mechanical strength and liability to damage due to the softness of the board.

The material used most extensively to date is p.t.f.e.-impregnated Fibreglass. As can be seen from Tables 2 and 3 the performance is quite close to that of pure polystyrene or p.t.f.e. P.T.F.E. in dispersion is applied to the Fibreglass mats, which are then piled and hot-pressed at an elevated temperature. No adhesive is used to bond the copper to the laminated sheet, which in part accounts for its good electrical properties. P.T.F.E.

Table 2

PROPERTIES OF MICROSTRIP LINES FOR DIFFERENT METHODS OF FABRICATION

Board	Strip conductor	Ground plane	Method of fabrication	Loss* dB/m at 4 000 Mc/s	Comment
Polystyrene	Cu foil	Cu foil	A	$<0\cdot3$	Brittle; low softening temperature.
Fibreglass-supported polystyrene (2 mats)	Cu foil	Cu foil	B	$\simeq 0\cdot3$	Low softening temperature.
P.T.F.E. impregnated Fibreglass	Cu foil	Cu foil	C	0·3	Best all-round material.
Expanded polyethylene	Cu foil	Cu foil	A	$\simeq 0\cdot6$	Flexible but easily dented; low softening temperature.
Polystyrene and one Fibreglass mat	Ag ink on Fibreglass mat	Cu gauze	D	0·9	Excellent adhesion; poor definition; low softening temperature.
P.T.F.E. impregnated Fibreglass	Ag ink	Cu foil	E	1·3	Poor adhesion; poor definition.
Polystyrene and 2 Fibreglass mats	Ag ink	Ag ink	F	1·3–2·3	Excellent adhesion; poor definition; low softening temperature.
Silicone-impregnated Fibreglass	Cu foil	Cu foil	C	2·2	Good mechanically; high water absorption.
Polyester-impregnated Fibreglass	Cu foil	Cu foil	C	2·9	As above.
As above	Al foil	Al foil	C	$\simeq 6\cdot0$	R.F. losses prohibitive.

Transmission line dimensions: $b = 5\cdot5$ mm, $h = 1\cdot5$ mm, $a = 25$ mm [see Fig. 2(b)].

Methods of fabrication:
- A Copper foil moulded to board and hand stripped.
- B Copper foil and one Fibreglass mat moulded each side; conductor hand stripped.
- C Commercially available copper-clad laminate; conductor hand stripped.
- D Conductor pre-printed on Fibreglass mat subsequently moulded to board, together with copper gauze on reverse side.
- E Hand-painted silver conductor on commercially available board.
- F As for D, but printed Fibreglass moulded both sides. Loss figures vary for different inks and method of consolidation.

* Measured by Deschamps method. Figures relate to loss at approximately 10 cm from transducer; asymptotic loss will be approximately twice the quoted figures.

Table 3

Summarized Properties of Some Copper-Clad Fibreglass Laminates

These materials vary considerably with the source of supply. Figures quoted must be regarded as typical only.

Resin	P.T.F.E.	Polystyrene*	Silicone	Epoxy
Maximum size sheet, in	37 × 16	—	37 × 16	37 × 16
Typical thickness tolerance,† in/in	0·006 5 on 0·032 : 0·007 5 on 0·063: 0·012 on 0·125			
Permittivity (dry)‡ ..	2·7	2·8	3·9	4·8
Loss tangent (dry at 1 Mc/s)	0·000 7	0·000 2§	0·001 5‖	0·018‖
Water absorption,** %	0·03	<0·05	0·30	0·24
Maximum continuous temperature, deg C	200	80	150	175
Approximate cost index relative to Phenolic grade	22		9	7

* Bond strength usually rather low.
† Tighter tolerances can generally be arranged.
‡ Electric field perpendicular to laminations.
§ Increased to 0·002 for high-impact polystyrene.
‖ Varies with choice of adhesive; may deteriorate by 10 times or more after immersion for 24 hours at 23° C.
** After immersion in distilled water at 23° C for 24 hours.

Fibreglass is in fact the best material available from the all-round point of view taking into account considerations of temperature, mechanical strength, water absorption, and freedom from warping. On the other hand, it is relatively expensive and is not available commercially in this country, although small quantities have been produced very successfully on an experimental basis.

In addition to the above there are a number of Fibreglass boards of intermediate cost and performance. These include epoxy-resin, polyester-resin, and silicone-resin boards. Of these silicone Fibreglass has found the most use so far; although not very suitable for microstrip it can be used quite successfully in high-Q triplate line.

It is important to stress that all Fibreglass boards suffer from certain disadvantages. First, the weave of the Fibreglass is pressed into the copper foil and results in a slightly corrugated surface. With insufficient care this can result in uneven etching and the production of pin-holes. Secondly, a Fibreglass board produced by one manufacturer may differ significantly from that produced by another, although using identical resins. This may occur because of different types of glass, different weaves and different ratio of cloth to glass. Where an adhesive is used this may vary also. Thus, whilst the two boards may be near enough identical in mechanical properties, water absorption, and even power factor, the difference in dielectric constant may be quite unacceptable in a microwave application. Finally there is the limitation that all Fibreglass boards (silicone in particular) deteriorate significantly after prolonged exposure to extreme climatic conditions such as may be found in the tropics.[34]

(5) MEASURING TECHNIQUES AND EQUIPMENT

(5.1) General Observations

Measurements at microwave frequencies may be divided in two classes; primary measurements providing a detailed description of the electromagnetic field, and secondary measurements concerned with the terminal pair properties, such as impedance or attenuation. The techniques used in determining the terminal pair properties of strip line or strip-line components do not differ appreciably from those used with other microwave transmission systems. In fact it has been the practice to make use, as far as possible, of existing equipment fabricated in coaxial line or hollow waveguide. This procedure is facilitated by the ease with which one can obtain a reasonably good match to coaxial cable over quite large bandwidths (see Section 6.1). Furthermore there is no reason to believe that the construction of precision instruments is any easier in strip line. In fact, apart from one or two possible exceptions, the difficulties would appear greater.

Where it is desired to investigate the actual field in the line there is then no alternative but to cut slots or insert probes. The results obtained by this procedure can be very misleading, however, particularly in the case of microstrip. An examination of the field in the air space above the strip conductor provides a very poor guide as to the field distribution in the region between the strip conductor and ground plane, where most of the energy is concentrated. On the other hand, an attempt to insert probes into the region between the strip conductor and ground plane can produce a serious perturbation of the field, and it is doubtful whether much significance can be attached to the quantities so measured.[19] For example, standing-wave measurements made in this way on a microstrip line suggest a considerable degree of mode interference, yet when the Deschamps technique[23] is used no evidence of spurious modes is found, provided that there are no serious discontinuities in the line structure.

(5.2) Measurement by Means of the Input-Impedance Diagram

As is well known the properties of a 2-terminal-pair network can be deduced from the measured short-circuit and open-circuit input impedances. At microwave frequencies it is more convenient to use a single termination which is moved through a distance of one quarter-wavelength in the transmission line terminating the network under investigation. The input impedance then describes a circular locus on the Smith chart concentric with the origin. This termination may be an open-circuit, a short-circuit, or a quite arbitrary lossy reactance—the generality of which incidentally is not widely appreciated.[20] With hollow waveguides satisfactory short-circuiting pistons can be constructed without much difficulty, thereby providing certain simplifications to the subsequent analysis. With strip transmission lines, however, a good movable short-circuit is not easily made, particularly when the space between strip conductor and ground plane is filled with dielectric. For this reason much of the earlier work on microstrip was conducted with an open-circuit termination, the position of which was moved by physically cutting the line. Owing to the small conductor spacing (1·5 mm) the radiation from an open-circuit is small and reflection coefficients of the order of 95% can be obtained. Unfortunately the method is destructive and precise repetition of a previous reading is usually impossible. In order to overcome this limitation the short-circuiting piston described in Section 5.3 was developed.

To measure the input impedance of the network a slotted section and standing-wave indicator may be used in the usual manner. Two examples of strip-line slotted sections have been described elsewhere,[6, 7] but in general their design, as has already been emphasized, presents a number of difficulties, particularly in the case of microstrip. For this reason it has been found preferable to connect the circuit under test to a conventional slotted line. For rough work the mismatch of the input transducer can generally be ignored. Where more precise results are required the transducer may be calibrated with the aid of the variable short-circuit described below. The information so obtained can be used to perform a relatively simple graphical

nalysis which gives both the scattering coefficients of the ransducer and the properties of the circuit under test. Further etails of this method, due initially to G. A. Deschamps, will be ound in References 20–23.

(5.3) Special Measuring Equipment

As mentioned in Section 5.2, it was found desirable to develop short-circuiting piston in microstrip. This device, illustrated n Fig. 6, comprises a grooved copper block which slides on a rass bed-plate to which is affixed a length of microstrip line. the block there is a bowed spring contact of beryllium-copper which presses lightly on the strip conductor. To absorb any energy which may travel on through the block and possibly resonate, the surfaces of the groove are coated with graphite and the microstrip line is terminated with a matched load. The load is also useful in that a matched-load termination can be obtained simply by removing the block. In order to appreciate how the device in fact operates it is useful to think of the piston as a short length of rectangular waveguide operating well below cut-off. From this it is evident that for maximum reflection the

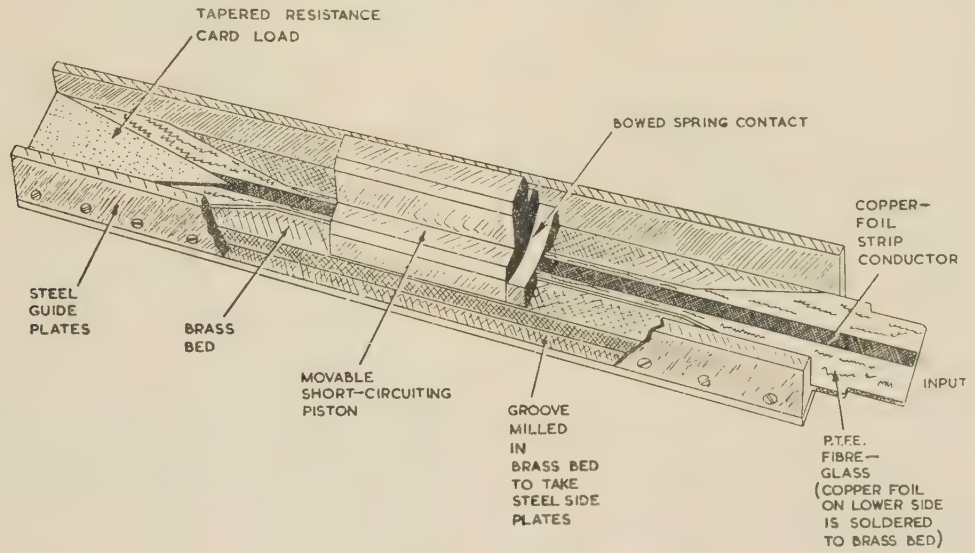


Fig. 6.—Microstrip short-circuiting piston.

o ensure a good contact between the sliding block and the bed-late, the two surfaces were separately ground and then lapped gainst each other. Steel side-pieces are fitted to stiffen the ed-plate and act as guides for the block, the under-side of which grooved to accommodate the microstrip line. The dimensions f the line are normal apart from the width of the Fibreglass neet. This is reduced to slightly less than twice the strip con-uctor width; any further decrease would produce a significant hange in the characteristic impedance. The upper surface of ne groove just clears the strip conductor at the front end of the lock and tapers upwards to the rear of the block to provide a nore substantial clearance of about 1 cm. At the front end of transverse dimension of the groove should be as small as possible. The measured results for the band 2 500–4 100 Mc/s were:

Average reflection coefficient: approximately 0·9.
Position of short-circuit: varied between 0·5 and 1·2 mm behind front edge.
Ratio λ_g/λ_0: constant at 0·69 within $\pm 0{\cdot}3\%$.

A second instrument of rather special interest is the wide-band attenuator illustrated in Figs. 7 and 8. The success obtained with the slotted-ground-plane experiments of Section 6.6 suggested the possibility of attenuators in which a slot of variable area was exposed to a lossy medium. Experiments showed the principle to be sound, provided that the power factor of the

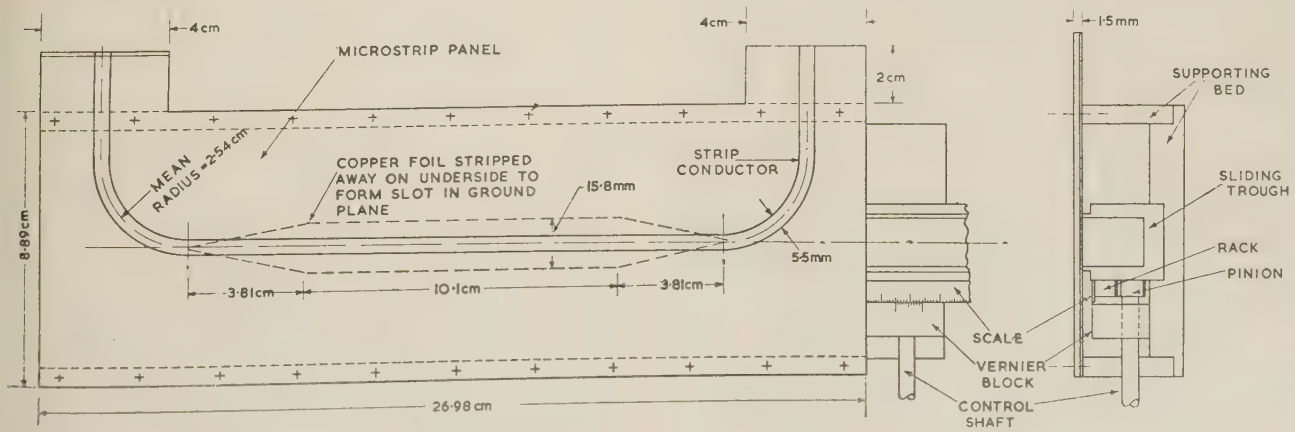


Fig. 7.—Microstrip laboratory attenuator.
General assembly.

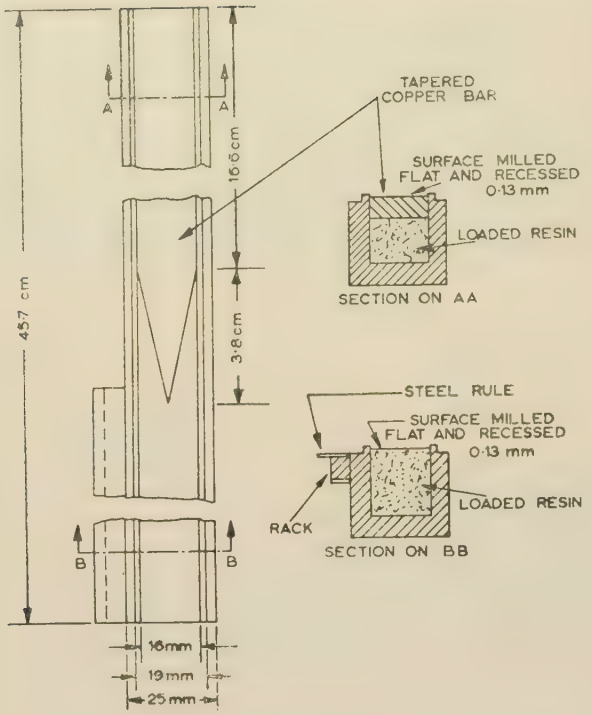


Fig. 8.—Microstrip laboratory attenuator.
Details of trough.

lossy medium was sufficiently high and that the depth was sufficient to ensure absorption without reflection.

The area of the slot may be varied in any one of several ways. However, the most obvious choice is to keep the transverse width of the slot constant, and vary the length in the direction of propagation. This is equivalent to a variable length of lossy line having a constant loss per unit length. It should, and in fact does, result in a calibration law which is substantially linear over most of the travel. The slot is tapered at each end so as to ensure a reasonable match, and the lossy material (50% Carbonyl 'E' iron dust in Marco resin, by volume) is contained in a sliding metal trough which is closed over approximately half its length by a tapered copper bar. Longitudinal movement of the trough effectively varies the length of the exposed slot.

To ensure a stable contact between the sides of the metal trough and the ground plane of the Microstrip line the following procedure was adopted. The top surface of each side wall of the trough is narrowed down and accurately milled flat. At the same time the microstrip line is supported in such a way that it is bowed downwards at the centre, and consequently forced against the top surfaces of the trough sides. The natural resilience of the Fibreglass is such that an adequate contact is made along the entire length without the aid of spring-finger contacts or similar devices. A final machining process on the trough ensures that the upper surface of the filling is a constant depth below the top surfaces of the trough sides. This is essential since the power absorbed varies with the proximity of the lossy medium.

The performance of this attenuator is illustrated in Fig. 9. If these results are re-plotted as a function of frequency it is found that in the linear range the incremental attenuation α, in decibels per centimetre travel, is given very closely by

$$\alpha = 0{\cdot}36 + 1{\cdot}21f$$

where f is in gigacycles per second.

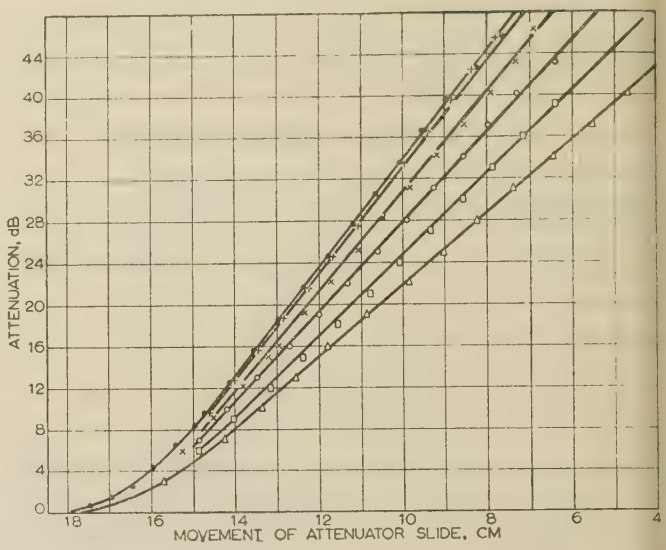


Fig. 9.—Calibration of microstrip laboratory attenuator.

● 4025 Mc/s	5·16 dB/cm.	○ 3240 Mc/s	4·32 dB/cm.
+ 3910 Mc/s	5·12 dB/cm.	□ 2905 Mc/s	3·92 dB/cm.
× 3600 Mc/s	4·76 dB/cm.	△ 2610 Mc/s	3·44 dB/cm.

Virtually no troubles have been experienced in using this attenuator. However, for attenuations in excess of 40 dB, the output side of the attenuator becomes a little sensitive to stray radiation or coupling from the input side. Although the effect is small, it means that a certain care is required in operating the instrument.

(6) COMPONENT DEVELOPMENT

(6.1) Range of Components

Quite a variety of components in strip-line form have been described in the recent literature. They include transducers, loads, attenuators, hybrid junctions, directional couplers, power-dividing networks, antenna systems and filters.* In addition, some work has been done on more specialized devices, such as ferrite attenuators and phase shifters and broad-band noise sources incorporating a gas tube.[11] In the following paragraphs, however, the discussion will be restricted principally to components not previously described elsewhere.

(6.2) Transducers

Transducers from waveguide or coaxial to strip-line may be of the in-line or, alternatively, the right-angle type. One example of each is illustrated in simplified form in Fig. 10 for the two principal strip transmission-line forms. For the triplate line the right-angle junction is not to be recommended, because it tends to stimulate an undesired parallel-plate mode between the two ground planes. However, in some cases it appears to have been used with success.[6] For microstrip the right-angle junction was found to give a better match over a larger bandwidth, and has since been universally adopted.[11] In an alternative arrangement which would appear equally satisfactory the coaxial line enters from above.

Since practical strip transmission lines are virtually non-dispersive over their useful operating range, the characteristic impedance is independent of frequency. The same is true of coaxial cables, hence transducers of the form shown in Fig. 10 are inherently wide-band, with no lower frequency limit. How-

* See *Transactions of the Institute of Radio Engineers*, 1955, Vol. MTT-3, No. 2 Special Issue 'Symposium on Microwave Strip Circuits'. Only the more relevant papers have been specifically quoted here.

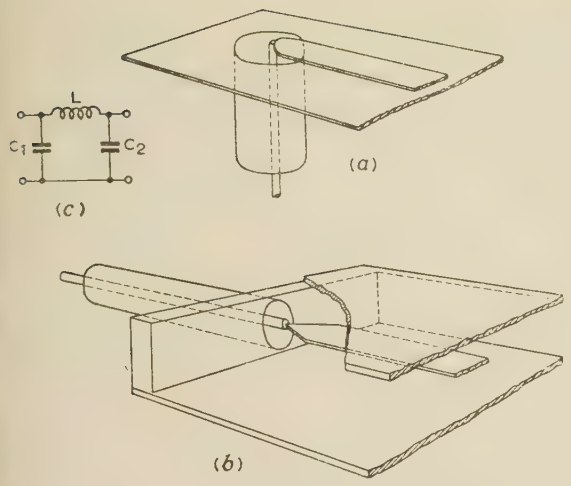


Fig. 10.—Coaxial to strip-line transitions.
(*a*) Right-angle junction (shown for line above ground).
(*b*) In-line junction (shown for triplate).
(*c*) Equivalent circuit, either junction.
Note: For simplicity supporting dielectric sheets are not shown.

wer, the physical discontinuity between the two systems introuces a reactive component which provides an upper frequency mit, but by good design this can be made to approach the pper useful frequency of the strip-line system, which, for easons discussed in Section 2.4, tends to lie for the moment in he region of 10–15 Gc/s. For example, with the right-angle nction of Fig. 10 the series inductance L of the protruding robe can be matched out over a very wide band of frequencies y a suitable adjustment of the capacitances C_1 and C_2, where C_1 is the discontinuity capacitance of the inlet hole in the ground lane and C_2 is the edge capacitance of the strip conductor. In ne of the examples quoted by Arditi[11] a match of better than ·7 is achieved over virtually the entire centimetric range with n upper limit of about 11 000 Mc/s. The same equivalent rcuit applies in the case of the triplate junction of Fig. 10(*b*); lthough the probe effect is smaller than in Fig. 10(*a*) there is ill an inductive component due to the perturbation of the agnetic field distribution in going from one geometry to nother.

The coupling between waveguide and strip line may be chieved in a variety of ways. In Fig. 11 is shown a transducer

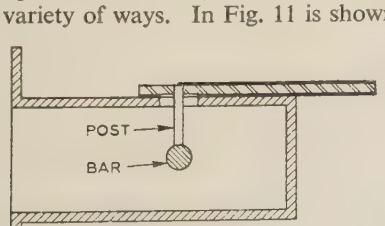


Fig. 11.—Waveguide to microstrip cross-bar transition.

hich is particularly simple to construct, in that use is made f an existing waveguide-to-coaxial cross-bar type transition. he voltage standing-wave ratio (v.s.w.r.) achievable is therefore f the same order as that of the original cross-bar transition on hich it is based. A similar form of construction may be used th triplate line, but in order to ensure symmetrical excitation preferable to mount the triplate line at right angles to the oad face of the guide; the post and the strip conductor are e in line. More recently a colleague, Mr. G. H. B. Thompson, a developed a tapered ridge-guide transducer outwardly similar to that described by Arditi.[11] The initial results have been very encouraging; for example, by a little adjustment, a match of 0·95 or better can be obtained over a 10% bandwidth.

(6.3) Loads and Attenuators

A satisfactory broad-band load, having a v.s.w.r. better than 0·9 over substantially the entire centimetric band, is easily made by attaching a tapered piece of carbon-backed card of prescribed shape to the strip conductor. If the position of the card can be adjusted so as to alter the amount of energy intercepted, a variable attenuator is obtained. This is usually done by sliding the card laterally across the line. However, where lateral space is at a premium a hinged-flap arrangement has been successfully used.

Carbon-backed card, although satisfactory for dummy loads, is by no means the ideal material for attenuators. First, its properties vary with humidity and temperature, and secondly, it is insufficiently rigid unless supported by a stiffer material. As an alternative, metallized-glass attenuator elements of the type used for hollow waveguide have been tried. Measurements on one such attenuator are shown in Fig. 12; the v.s.w.r. for

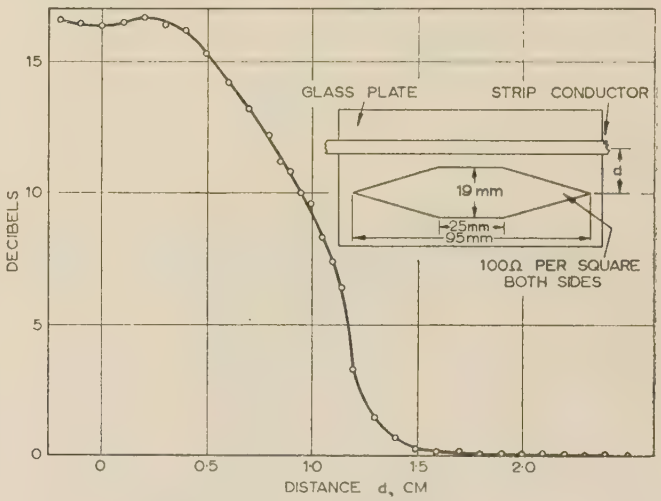


Fig. 12.—Glass slide attenuator.
$f = 4\,000$ Mc/s.

maximum attenuation was 0·8. For reproducible results the slide must remain at a fixed distance above the strip conductor, and for maximum attenuation the slide must actually touch the strip conductor. (This is the condition under which the measurements illustrated were taken.) Mechanically, however, this is not satisfactory since the continual abrasion of the metal film eventually disturbs the performance of the attenuator.

As an alternative, the metallized-glass element was placed between the strip conductor and the ground plane. The maximum attenuation was then only 5 dB. This, and the fact that in Fig. 12 the maximum attenuation occurs when the element is slightly off-centre, confirm that the surface attenuators operate by absorption of fringe field energy. At the sides of the conductor the E lines are roughly horizontal and therefore in the plane of the absorbing surface.

Attenuators and loads involving insertion of lossy material between strip conductor and ground plane will work satisfactorily only if the material is homogeneous and extends the full distance between ground plane and strip conductor.[24] This type of attenuator, however, suffers from a number of mechanical difficulties.

Because of these various difficulties it is considered that, where accurate and reproducible results are required, an attenuator on the lines of that described in Section 5.3 is likely to prove more satisfactory.

(6.4) Crystal Mounts

Conventional crystal capsules are for the most part designed for inclusion in a coaxial mount. With one exception,[19] therefore, strip-line crystal mounts have consisted in effect of a coaxial mount coupled by a suitable transducer (see Section 6.3) to the strip-line system. Hence, the only advantage that such mounts have over conventional coaxial is that any matching elements required may be incorporated in the strip-line system rather than the coaxial system. Fig. 13 shows a mount which

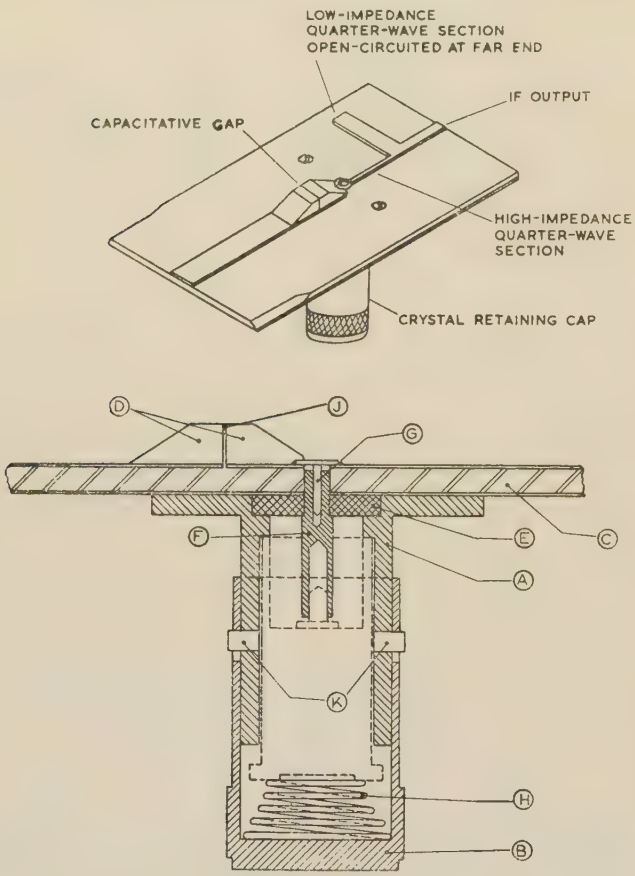


Fig. 13.—Wide-band crystal mount.

A Crystal-holder body.
B Crystal retaining cap.
C Microstrip panel.
D Capacitive lugs.
E Styrene washer.
F Centre pin.
G Centre-pin retaining screw.
H Crystal retaining spring.
J Mica spacer.
K Bayonet fixing lugs.

exploits these possibilities very fully. This circuit differs from the more conventional arrangement in that the crystal is in direct shunt across the transmission line. It is therefore necessary to provide a series isolating condenser on the r.f. side, having a low impedance at radio frequencies and a high impedance at intermediate frequencies. This arrangement has the advantage that the i.f. output can be taken directly from the strip conductor rather than the metal cartridge of the crystal, thus allowing easy removal of the crystal. Furthermore, the output filter can be printed on the surface of the panel instead of fabricated as part of the coaxial mount.

In practice it is convenient to make the series condenser fulfil the double function of isolation and r.f. matching. This is achieved by placing it at a point in the line where the crysta impedance appears inductive. Since the crystal presents a poo match to the line quite a small capacitance suffices. In th particular example illustrated, which is designed for the banc 2 500–4 100 Mc/s, a capacitance of 2–3 pF is required. Experi ments were conducted at 4 000 Mc/s to see whether this could b achieved by a simple lateral gap in the strip conductor. Fo gaps in the measured range 0·005 in to 0·050 in, the relationshi between the normalized series reactance X and the gap width d (in inches) was given quite closely by a linear law:

$$X = 2{\cdot}68 + 146d$$

This implies a series capacitance of only 0·33 pF for a gap o 0·005 in, which is regarded as the limiting size in a practica application. Instead, a small brass block was soldered to the strip conductor through which a 0·010 in slot was milled at ar appropriate distance from the crystal connection. From th production point of view it would no doubt be preferable t obtain the desired capacitance by a convoluted slit, and it i understood that this can be done.[29]

This arrangement leads to a very low shunt capacitanc across the i.f. terminals, in this case only 6 pF, and the mount i therefore suitable for the detection of very wide-band signals, the upper frequency limit being approximately 200 Mc/s. This is due partly to the low value of input series capacitance, and partly to the manner in which the i.f. connection is made. Since the output filter is printed, a relatively thin series conductor can be used having a high transmission-line impedance of the orde of 150 ohms. This in turn allows the use of a relatively high impedance (of the order of 60 ohms) for the terminating shunt section. A further advantage of the arrangement is the eli mination of a d.c. return, the precise position of which is generally rather critical.

Without the series capacitance the input v.s.w.r. of the mount averages about 0·2 over most of the band. With the capacitance the match is peaked to about 0·8 at 3 420 Mc/s, falling away gradually to the original value of 0·2 at the extreme frequencies of 2 500 and 4 100 Mc/s. In these tests a crystal current of 1 mA and d.c. load of 100 ohms were used. A somewhat better match can be obtained by placing an inductive tuning element inside the cartridge close to the crystal itself, but the additional mechanical complication is not worth the small improvement achieved. Although the match of these crystals is poor they are very consistent, as can be seen from Table 4. This is of con-

Table 4

DISPERSION OF CRYSTAL-HOLDER MATCH

Frequency, Mc/s	2 600	3 400	4 100
Seven CV 2154 and CV 2155 crystals tested in one holder	1·5%	2·3%	2·2%
One CV 2154 crystal in 5 identical mounts	2·9%	1·4%	1·5%

The figures given represent r.m.s. deviation of v.s.w.r. as a percentage of the averag v.s.w.r. at each frequency. The dispersion of reflection-coefficient angle is of the sam order.

siderable advantage in the design of balanced mixers. Table 4 also shows the reproducibility of five identical microstrip holders Owing to the simplicity of the design and the inherent repro ducibility of the printed-circuit technique this is of a high order

(6.5) Hybrid Junctions

The conventional hybrid-T circuit employing congruent serie and shunt junctions cannot conveniently be realized in strip-lin

rm. However, an equivalent performance can be achieved y other means. In the first place, there exists a variety of ring- pe circuits which will provide the desired property of con- gacy through the use of interconnected quarter-wave trans- ission lines. These are easily realized in strip-line form, the nly limitation being the inherent frequency sensitivity of the uarter-wave sections. Secondly, there are various types of 3 dB irectional coupler, a particular example of which is described detail in Reference 25. Where the coupling mechanism in e directional coupler consists of quarter-wave sections the istinction between the second two categories breaks down, and is therefore convenient to deal with such circuits in the present ection under the general heading of ring circuits.

The best known of these ring circuits is the so-called 'rat-race' f Fig. 14(*a*). A number of these junctions in microstrip form

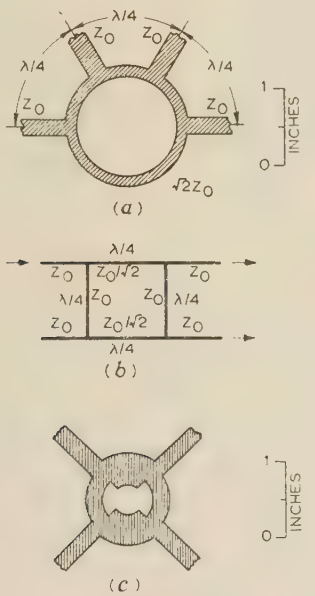


Fig. 14.—Simple hybrid junctions.

(*a*) Conventional rat-race.
(*b*) Schematic of 3 dB branched-arm coupler.
(*c*) 3 dB branched-arm coupler rearranged.

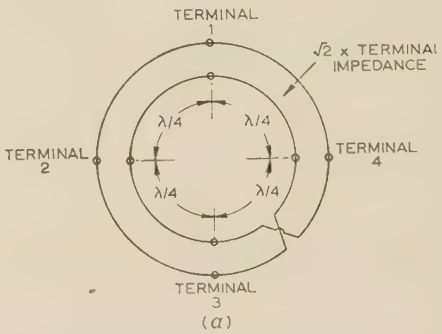
Note: Scales relate to a design frequency of 3 200 Mc/s, a 50-ohm line impedance, nd a sheet thickness of 1·5 mm.

ave been built and tested, and it is usual to obtain a decoupling n excess of 20 dB with a power split equal within 1 dB over andwidths of about 20%.[11] For narrow bandwidths a ecoupling in excess of 30 dB is possible, provided that the omponent is fabricated with sufficient precision. As an alter- ative, the 3 dB branched-arm coupler of Fig. 14(*b*) may be used; wing to the elimination of the $\frac{3}{4}\lambda$ segment a somewhat reduced requency sensitivity is to be expected, and this has been confirmed xperimentally. Because of the extreme width of the low- mpedance conductors the exact electrical length of the branch rms becomes rather uncertain, and for this reason it has been ound advantageous to rearrange the circuit as shown in Fig. 14(*c*). specimen having a design centre frequency of 3 300 Mc/s was ested and gave a power split equal within 1 dB from 2 700 to 00 Mc/s, which is over twice the bandwidth usually possible ith a rat-race for the same performance. There was no evidence f any improvement in decoupling, however, which was down 0 dB at the extreme frequencies.

It should be noted that Figs. 20(*a*) and (*c*) are drawn to the ame scale and are based on the same design centre frequency f 3 300 Mc/s, a line impedance of 50 ohms, and dielectric-sheet hickness of $\frac{1}{16}$ in. If the design centre frequency were to be increased, all other parameters remaining unchanged, the inner circumference of the ring would eventually decrease to zero. There is therefore an upper frequency limit to these ring circuits which is reached more quickly for the 3 dB branched-arm coupler than for the rat-race. This can be overcome by increasing the line impedance or by decreasing the thickness of the dielectric sheet, the latter solution being generally preferred provided that the small increase in attenuation can be tolerated. A further point to be noted is that when the inner diameter becomes small the effective electrical length of the ring is no longer equal to the mean circumference and a small correction is necessary.

In order to increase the operating bandwidth of the branched-arm coupler the number of branches may be increased to three or more. The impedance of the outer branches tends to become impracticably high, however, and it has therefore been found preferable to resort to the slot-coupling technique described in Reference 25.

A very considerable improvement to the conventional rat-race can be obtained by replacing the $\frac{3}{4}\lambda$ segment by a $\frac{1}{4}\lambda$ segment in conjunction with a phase reversal. This arrangement, known as the reverse-phase hybrid ring,[33] is illustrated schematically in Fig. 15. It is possible to show that this circuit



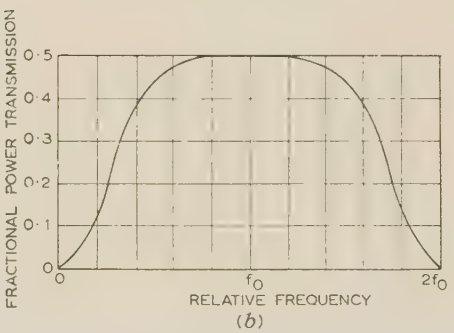


Fig. 15.—Reverse-phase hybrid ring.

(*a*) Schematic.
(*b*) Theoretical power-transmission curve.

ensures perfect decoupling and power division over an infinitely wide frequency band. The input match is still a function of frequency, however, and as the frequency of operation departs from the design centre frequency an increasing amount of power is lost by reflection. An example of an experimental ring circuit executed partly in microstrip and partly in balanced strip line is illustrated in Fig. 16. The phase reversal which is situated in the part of the ring circuit executed in balanced line is illustrated in greater detail in Fig. 17. The discontinuity susceptance of the phase reversal is quite small, and there is no significant evidence of radiation. The performance of the ring circuit as a whole is illustrated in Fig. 18.

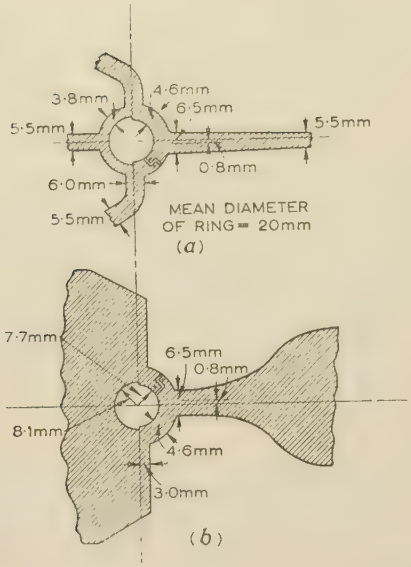


Fig. 16.—Experimental reverse-phase hybrid ring.
(*a*) Upper conducting system.
(*b*) Lower conducting system.
Printed on 1·5 mm p.t.f.e. Fibreglass.

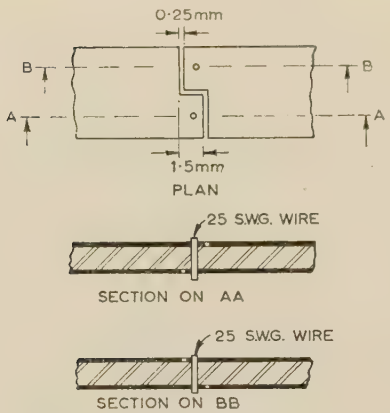


Fig. 17.—Details of phase reversal.

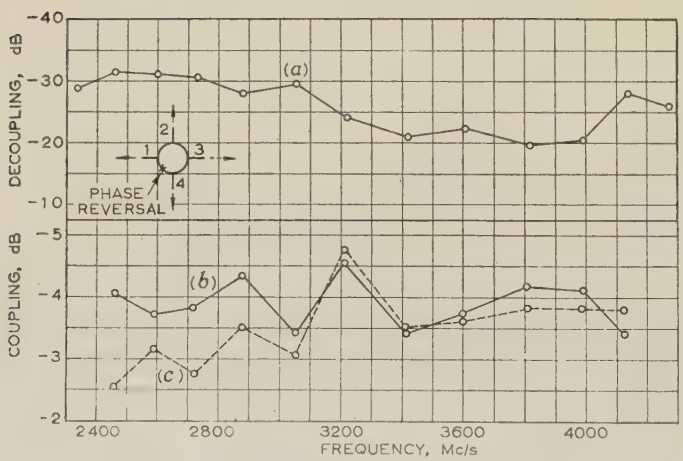


Fig. 18.—Experimental reverse phase hybrid ring.
Measured results.
(*a*) 20 log $\left|\frac{E_3}{E_1}\right|$. (*b*) 20 log $\left|\frac{E_2}{E_1}\right|$. (*c*) 20 log $\left|\frac{E_4}{E_1}\right|$.

A balanced mixer based on the above hybrid ring, together with two crystal mounts as described in Section 6.4, is illustrated in Fig. 19. In order to simplify the mounting of the coaxial

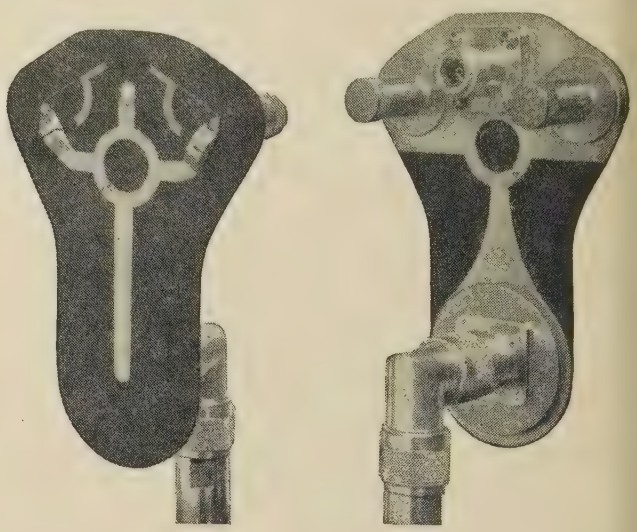


Fig. 19.—Balanced mixer employing reverse phase hybrid ring.

connector of the lower input arm, the lower strip conductor is flared out to form a ground plane for the upper strip conductor. However, there exists a danger that small currents flowing on the underside of this conductor may find their way along the outside of the coaxial cable and eventually reach the main ground plane via the outside of the other coaxial outlet, or via the crystal holders. For this reason a planar radial choke is provided.

In conclusion, it may be of interest to mention the novel 5-arm junction illustrated in Fig. 20. This component is of

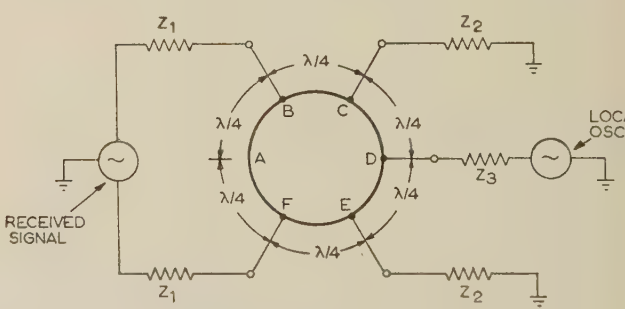


Fig. 20.—Six-arm hybrid junction.

value where the receiving antenna is a dipole, since it avoids the use of a balance-to-unbalance transformer. Basically five arms are provided, and these are connected to the external circuit in the manner shown. The circuit provides local-oscillator noise cancellation, and decoupling of the local oscillator and antenna circuits. The necessary matching conditions may be deduced by splitting the ring circuit on the axis of symmetry passing through the points A and D. They are achieved when

$$Z_2 \times 2Z_3 = Z_{CD}^2$$

and

$$Z_1 \times Z_2 = Z_{BC}^2$$

where Z_1, Z_2 and Z_3 are the external circuit impedances, and Z_{BC} and Z_{CD} are the characteristic impedances of the transmission lines BC and CD. It will be noted that the impedance

of the line BAF remains unspecified. In practice an additional arm is added to the junction at A and provided with a short-circuiting post at a distance of one quarter-wave from A, thus providing a common d.c. return for the two crystals.

(6.6) Directional Couplers

Directional couplers of the branched-arm type have been mentioned briefly in the last Section. For narrow bandwidths and with large coupling they are no doubt quite satisfactory, but otherwise there is a tendency for the characteristic impedance of the branch arms to assume impracticably high values, e.g. greater than 150 ohms. For coupling values less than −10 dB a more practical proposition is the parallel-line coupler. Two strip conductors situated in close proximity are printed on the same board. Coupling occurs through fringe-field interaction, and because the lateral attenuation of the field is rapid the conductors must be closely spaced. There are two principal limitations to this device, first, the coupling varies rather rapidly with frequency, and this variation cannot conveniently be controlled; secondly, there is evidence that the coupled section can resonate in a balanced twin-strip mode. The latter tends to produce erratic variations in directivity similar to the long-slot effect in slot-coupled waveguides. In an attempt to overcome the various limitations of the previous designs a type of directional coupler has been developed comprising two microstrip transmission lines mounted back to back and coupled through a series of discrete slots in the common ground plane. By a suitable choice as to size, position and number of the coupling holes a wide variation in coupling strength is possible. Moreover, according to the shape of the hole, either forward or reverse directivity can be obtained. Finally, by modifying the phase velocity in one of the two lines the coupler can be broadbanded so as to give virtually constant coupling over bandwidths as large as 1·6 : 1. This work is described in a companion paper.[25]

(6.7) Power-dividing Networks

A potentially fruitful application of strip-transmission-line technique occurs in the design of power-dividing circuits, where a large number of separate outputs are to be derived from a single input.[4, 26] Usually these networks are built up in the form of a tree from one or other of the simple two-way dividers shown in Fig. 21. In the first divider matching is achieved by the use of quarter-wave transformers, whereas in the second a gradual taper is used.

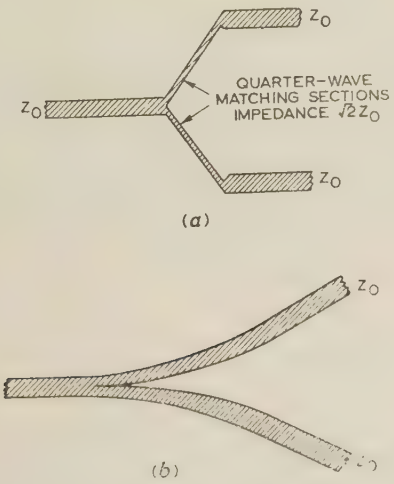


Fig. 21.—Binary power dividers.
(*a*) With quarter-wave transformers.
(*b*) Tapered.

The successful design of the tapered dividing section depends on a proper appreciation of the fringe-field interaction at the point of division. This is illustrated in Fig. 22 for a worked

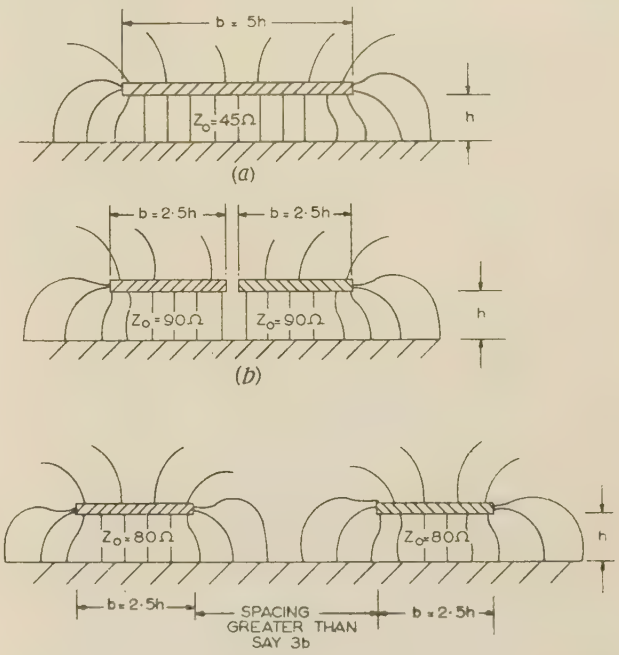


Fig. 22.—Fringe-field effects with power dividers.
(*a*) In contact. (*b*) Contiguous. (*c*) Remote.

example comprising, for simplicity, an air-spaced line. At the point of bisection, Fig. 22(*b*), the effective impedance of each line referred to the generator is evidently twice that of the unbisected line. Looked at another way, the one conductor in effect acts as a guard conductor for the other. However, when the lines have separated a sufficient distance, as shown in Fig. 22(*c*), the fringe fields take up their normal condition on both sides of each conductor, and hence the impedance falls from 90 ohms to about 80 ohms. From this illustration it can be seen that for low reflection the conductors must separate very gradually. Alternatively the input line should split into two separate conductors which diverge as rapidly as possible, the width of each being chosen to give 90 ohms when considered separately.[26] It is believed that the former method gives the smoother transition.

With these principles in mind a 4-way power split with gradual tapers was built. In this design the strip conductor gradually widens so as to reduce the impedance to about 20 ohms, at which point the conductor divides. The separation between the two conductors is gradually increased until the mutual interference of the fringe fields can be neglected, at which point the two conductors divide again, the impedance level at this point being about 30 ohms. This gives a very smooth transition, with a high degree of symmetry between the four outlets and an overall length of only 12 cm. Over the frequency range 2 500–4 100 Mc/s the average input v.s.w.r. was 0·85, and the equality of power division was better than $\frac{3}{4}$ dB at most frequencies.

(6.8) Filters

(6.8.1) General Considerations.

Basically there is no difference in technique between strip-line and other types of transmission-line filter. However, if the

fullest possible advantage is to be taken of printed-circuit techniques all the elements should be in planar form. In itself this is not a serious restriction, for there exists an adequate range of such elements. On the other hand, since the designer has only two dimensions to play with, he is frequently restricted in regard to the range of electrical values possible. An example of this, which has already been mentioned, is the series capacitor used on the broad-band crystal mount (Section 6.4), and a similar difficulty occurs with shunt capacitors, since the strip conductor frequently tends to assume dimensions which are prohibitively large fractions of a wavelength. This can, of course, be overcome by a local reduction of the ground-plane spacing, but mechanical complications of this kind are undesirable.

These difficulties are particularly evident when designing filters with constant-k or m-derived sections. To obtain the precise cut-off frequency desired a fair amount of experimental adjustment has been found necessary. Furthermore, it is frequently impossible to ensure freedom from spurious responses over a sufficiently wide bandwidth. Despite these difficulties a very considerable measure of success is understood to have been achieved with these types of filter in applications where the operational requirements would appear not to be too stringent.[5, 27–29]

Another important consideration is the Q-factor of the elements, and in this respect high-Q triplate line has a definite superiority. In the first place the dissipative losses are a good deal lower, and secondly, it is easier to limit the loss by radiation or mode conversion (see Section 2.3). This does not exclude the possibility of satisfactory filters in microstrip, however. In the case of low-pass filters it is not essential for the elements to have a very high Q-factor, and it is sufficient to ensure the selection of suitable elements noteworthy for the absence of any marked tendency to radiate. Alternatively, as in the case of band-stop filters, it has been found that the problem may be circumvented quite successfully by fabricating the critical elements as enclosed cavities which are then coupled to the microstrip line by suitable means.

(6.8.2) The Tapered-Re-entrant-Line Low-pass Filter.

In an attempt to overcome the various limitations mentioned above a new type of low-pass filter has been developed. It operates by virtue of recurrent wave interference between two transmission lines having in effect different velocities of propagation, and is a development of the re-entrant transmission-line filter originally described by Andrew Alford[30] in 1938. Fig. 23

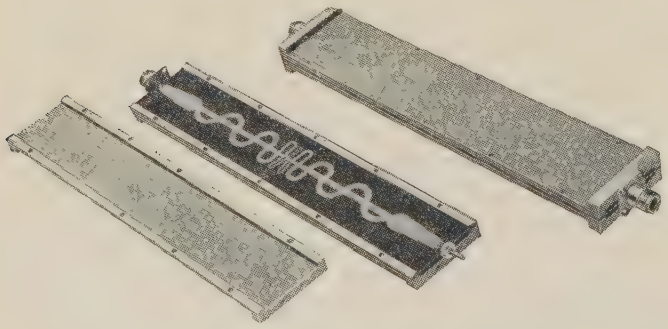


Fig. 23.—Tapered re-entrant line low-pass filter.

shows the strip conductor configuration and general construction of one of these filters, and Fig. 24 a typical response curve. The rate of rise at the cut-off point is extremely rapid, namely 50 dB at about 100 Mc/s. A small dip in attenuation to 28 dB occurs at 8 600 Mc/s, but otherwise the insertion loss is greater than

Reproducibility of four identical filters

Frequency	Deviation of insertion loss from mean			
	Filter 1	Filter 2	Filter 3	Filter 4
Mc/s	dB	dB	dB	dB
2465	−0·17	+0·02	+0·02	+0·12
2755	−0·16	−0·16	+0·06	+0·27
3070	−0·10	−0·10	+0·14	+0·06
3455	+0·26	−0·101	−0·06	−0·10
3655	0	0	0	0
3862	+0·07	−0·126	+0·03	+0·025
3980	+0·33	−0·039	−0·14	−0·14
4040	−0·32	−0·11	+0·11	+0·32
4135	−0·18	+0·08	−0·08	+0·19
4140	+1·86	+0·35	−1·11	−1·11
4240	−1·64	+0·62	+0·29	+0·73

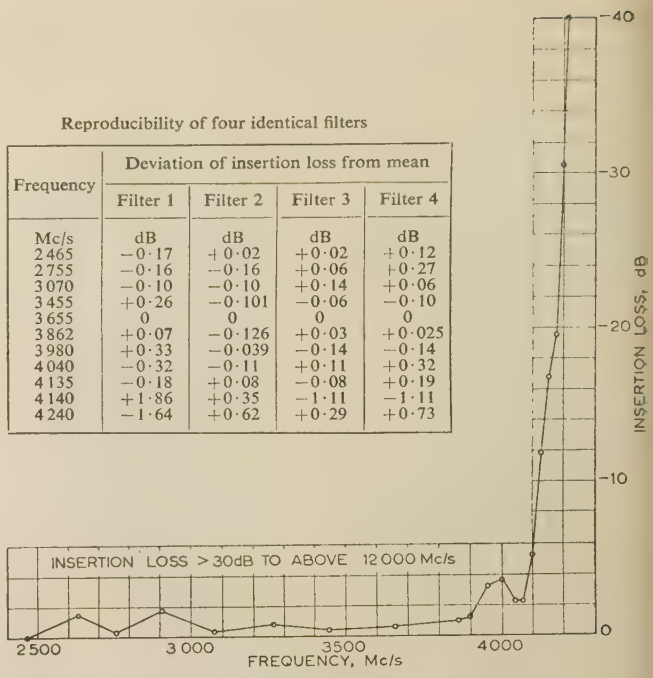


Fig. 24.—Tapered re-entrant line low-pass filter.
Measured results.

30 dB at all frequencies up to the highest measured frequency of 12 000 Mc/s. A particular feature is the high repeatability in the pass band; as can be seen from Fig. 24, the spread between four identical filters is less than 0·5 dB right up to the cut-off point. No special techniques were used to ensure this high repeatability, apart from accurate machining of the filter box and the use of a set of special dowel pins to ensure accurate registration of the prints on the two sides of the Fibreglass board.

It will be noted that, in the example shown, the high-Q triplate form of construction has been adopted. Filters of the same type have been constructed in microstrip and equally rapid rates of rise at the cut-off frequency have been obtained, as also equivalent insertion-loss characteristics in the pass band. However, a good deal of difficulty was experienced in screening the microstrip filter, since the natural resonances of the screening box tend to be excited, thus producing a perturbation of the response characteristic.

The theoretical principles involved in the design of these filters cannot be dealt with properly here and they will therefore be described in a later paper. It is worth pointing out, however, that although the theory of these filters is somewhat difficult their physical realization is extremely straightforward. For example, the measured cut-off frequency of the first model was within 1% of the theoretical, and consequently never needed any subsequent readjustment. The principal disadvantage is that, per decibel attenuation, the overall length is somewhat greater than for the conventional constant-k type filter.

Filters of this kind with higher cut-off frequencies, in the range 7 000–8 000 Mc/s, have also been made and have performed with equal satisfaction.

(6.9) Miscellaneous Components

Certain components have necessarily been excluded in the interest of brevity. Two recent developments should be mentioned in passing, however. One is a new form of high-pass filter capable of satisfactory operation up to 3 or 4 times the cut-off

requency. This, in combination with the low-pass filter escribed above, has allowed the design of wide-band band-ass filters. The other is a novel form of microscrip isolator of xtremely compact and simple form. Details of this latter evelopment were given by L. Lewin at The Institution's Con-ention on Ferrites, 1956.*

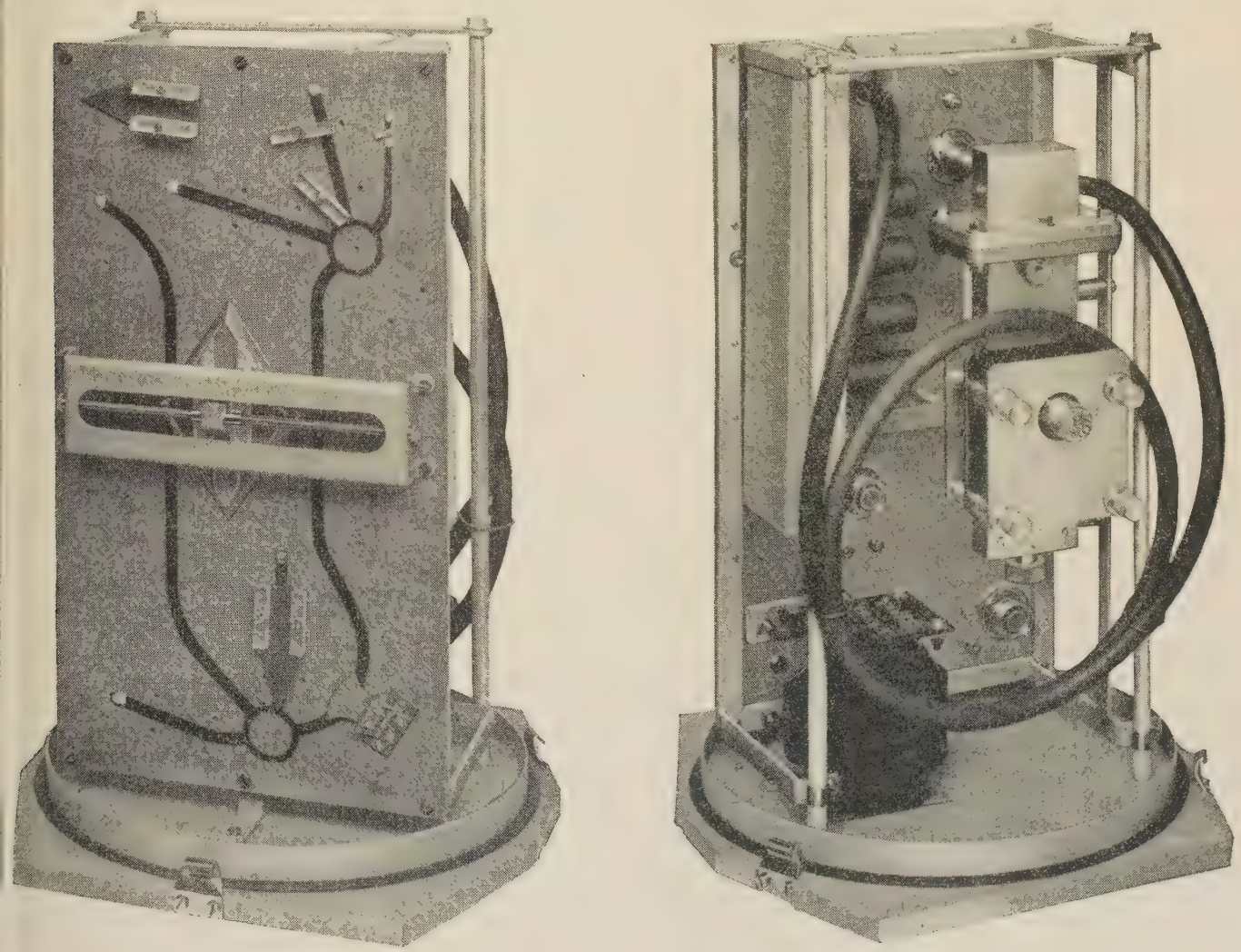
7) DUPLEX-WORKING TRANSMITTER-RECEIVER FOR THE FREQUENCY RANGE 3 800–4 200 Mc/s

(7.1) Introduction

A description of this unit is included in the present paper as provides a good illustration of the way in which isolated micro-trip components may be integrated to form a single assembly. The unit, illustrated in Fig. 25, forms part of an equipment used for the routine testing of the propagation properties of a micro-wave path. These measurements involve transmission in one direction only, but for the purpose of lining up the antennae and for other similar adjustments it is useful to have 'talk back' facilities in the reverse direction. Fig. 26 shows the composition of the r.f. head; it will be observed that the same oscillator is used for both transmission and reception. For transmitter-receiver No. 2 the frequencies f_1 and f_2 in Fig. 26 are interchanged and the antenna is connected instead to terminal B.

Equipment for this purpose incorporating standard waveguide has of course been available and in use for some time. The new unit, however, has certain advantages. The volume is about two-thirds that of the original unit; a worthwhile reduction in weight has been achieved; the unit can be set up more quickly to any desired pair of frequencies 60 Mc/s apart in the band 3 800–4 200 Mc/s. Furthermore, a change to other frequency spacings is relatively simple, and the adjustment of leak-through power from the master oscillator to the mixer is largely independent of the frequency-setting control. The performance is comparable with that of the standard waveguide head, except for a slight increase in overall antenna/crystal-mixer insertion loss of the order of 1·5 dB.

The microstrip panel is fabricated from $\frac{1}{8}$ in copper-clad p.t.f.e. impregnated Fibreglass and, unlike the previously described components, employs a transmission-line impedance of 70 ohms. For rigidity the panel is bolted to a backing plate of $\frac{1}{16}$ in aluminium.



(*a*) (*b*)

Fig. 25.—Microstrip duplex working r.f. head.

* LEWIN, L.: 'A Resonance Absorption Isolator in Microstrip for 4 Gc/s', *Proceedings I.E.E.*, 1957, **104** B, Suppl. 6, p. 364.

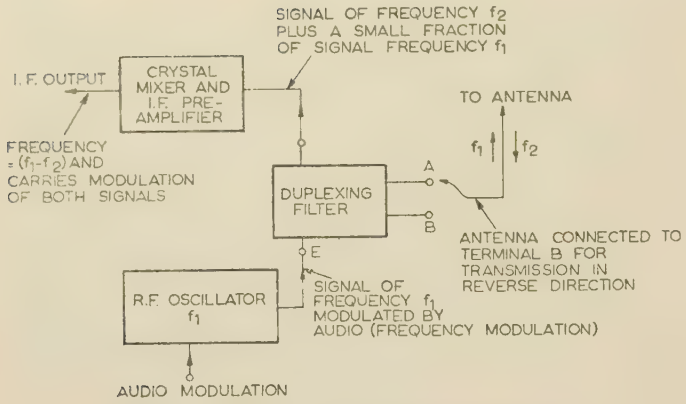


Fig. 26.—Block diagram of duplexing system.

(7.2) Detailed Description of the Circuit

A detailed schematic of the microstrip filter panel is given in Fig. 27. The r.f. oscillator (mounted at the rear of the panel) generates a signal of frequency f_1, which splits equally between the arms P and N of rat-race No. 2, the signal at P being in phase with that at N. Only a very small amount of power leaves by the arm Q, but it is just sufficient to ensure near optimum operation of the crystal mixer circuit. The length of the transmission line PDCK relative to the transmission line NM is such that the signals of frequency f_1 appearing at the points K and M are 180° out of phase. In this case all the power of frequency f_1 flows out through the arm J into the antenna circuit, and only a negligible quantity is dissipated in the matched load.

At the same time a signal of frequency f_2 is received by the antenna and splits equally between the arms K and M of the rat-race No. 1, the signal at M being 180° out of phase relative to the signal at K. Only a negligible quantity of power is dissipated in the matched load in the arm L. The length of the transmission line PDCK relative to the transmission line NM is such that the signals of frequency f_2 appearing at the points P and N are 180° out of phase. In this case the signals combine in the arm Q and cancel in the arm O. Hence, virtually all the received power is delivered to the crystal mixer circuits, and none is dissipated in the local oscillator.

The requirements with regard to the lengths of the transmission lines PDCK and JMN are given quite simply by the relationship

$$n = \frac{f_2}{2(f_2 - f_1)}$$

where n is the difference in length between the two lines expressed as the number of electrical wavelengths at frequency f_1, and f_2 is assumed larger than f_1. However, in deducing this relationship it is assumed that n is an integer. For this reason it is not possible to make an absolutely independent choice of f_1 and f_2. Nevertheless, it is possible to approximate the desired conditions very closely, and the bandwidth of the filter is sufficient for any error caused to be of negligible importance.

In the present case the design specification required f_1 and f_2 to be 3790 and 4030 Mc/s, respectively. From eqn. (1), $n = 33{\cdot}6$, the nearest integer being $n = 34$. The precise values of f_1 and f_2 are thus 3970·75 and 4030·00 Mc/s. Allowing for the lengths of MN, KC and DF, this gives, in practice, a cable of approximately 1·6 metres in length.

As the frequency is gradually raised from f_1 to f_2, the output power gradually shifts from arm J to arm L. Hence, for the complementary equipment at the second terminal station the antenna is connected to terminal B instead of terminal A. Likewise the matched load is transferred from arm L to arm J. Otherwise the two equipments are identically tuned.

It can be shown that the response of this filter is sinusoidal with frequency, thus providing a fairly broad region over which operation is possible without any significant variation in loss. This property is clearly exhibited in the measured results of Fig. 28. A second property of importance is that under ideal conditions the filter presents a perfect match at both its input and its output terminals for all values of f. In practice, owing to the number of junctions, the input match to the unit as a whole varies from about 0·5 to 0·8 over the operating band 3800–4200 Mc/s.

In order to equalize the loss in the short and the long transmission lines a small attenuator is provided in the path NM. By this means it is possible to ensure that at the frequency f_2 the power delivered to the arm O is very small indeed, e.g. −30 dB relative to the arm Q. The degree of rejection provided is, in fact, in excess of the operational requirements. To cover small changes in frequency a special tuning control is provided. This device, which is a form of phase-shifter, or line-stretcher, comprises a tapered dielectric slab which may be slid over the microstrip line, thus modifying the velocity of propagation in the line. To provide both positive and negative variations a circuit configuration has been chosen which allows the slab to be moved either over the line KC or over the line MN. The slab, which measures approximately $4\frac{3}{8}$ in and 3 in between opposite corners, is fabricated from styrene and provides a tuning range of ±17 Mc/s. Operation of this control shifts both resonant

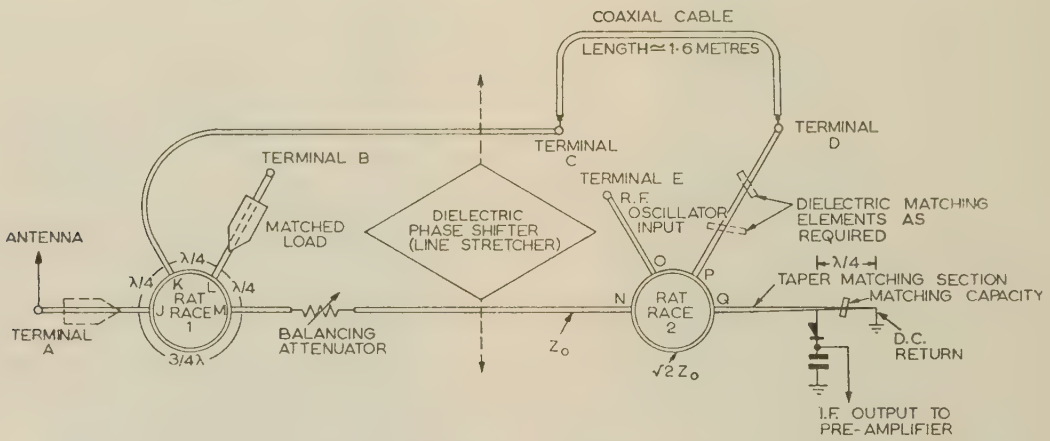


Fig. 27.—Schematic of duplexing filter.

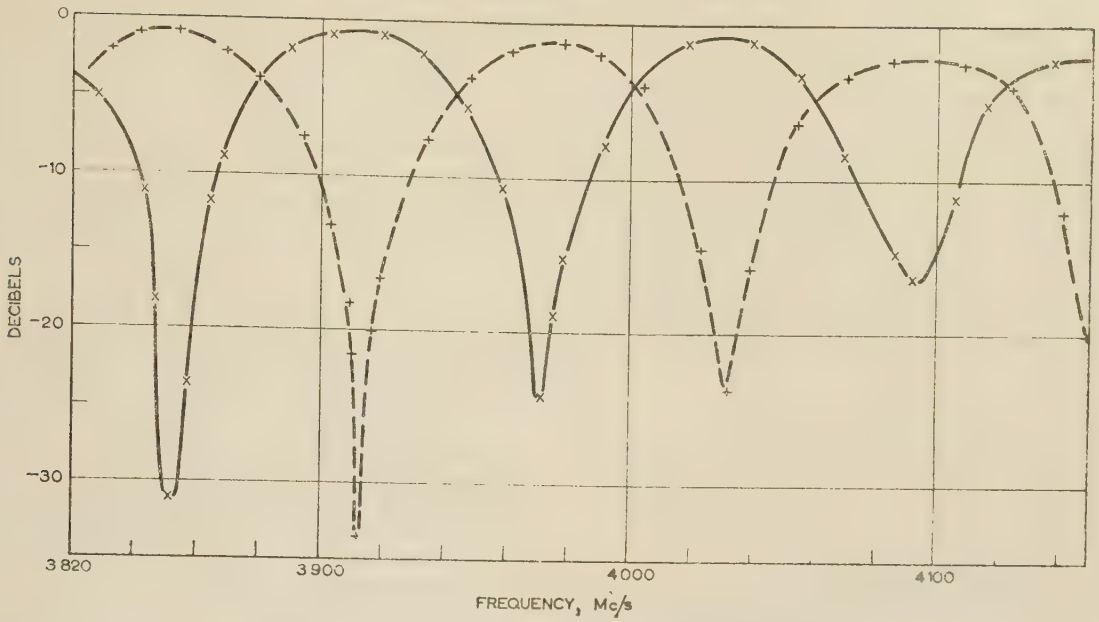


Fig. 28.—Insertion loss characteristic of duplexing filter.
—— Input A. – – – Input B.
Output monitored by crystal mixer.

equencies simultaneously, and leaves the frequency spacing $-f_1$ virtually unaltered.

A number of matched loads with styrene clamping bars are rovided for lining up the unit. Each load consists of a carbon-acked card 2½ in long by 1 in wide having a surface resistivity f 100 ohms. Only one is required for normal operation and the emainder may conveniently be stacked in a spare position at the op left-hand corner of the panel (see Fig. 25).

The requirement that the leakage of oscillator power from he arm O to the arm Q be controllable over a range of about 23 to −30 dB requires an additional form of tuning control. his is so because quite small mismatches at the arms P and N f rat-race No. 2 will cause a considerable variation in decoupling etween arms O and Q. It was found experimentally that the yrene slabs used for clamping the loads in position could be sed as tuning elements capable of giving the small amount of orrection required. These slabs may be placed as shown in ig. 25 so as to form an elementary double-stub tuner; the ositions are adjusted empirically until at the frequency f_1 the ght amount of crystal current for optimum mixer operation is chieved.

No special difficulties were encountered during testing and the nit handled easily. However, the change from the more usual alue of 50 to 75 ohms has the disadvantage of increasing the elative amount of power carried in the air space above the ielectric. In a balance condition the placing of metallic objects ithin about one inch of rat-race No. 2 will produce a noticeable teration in oscillator leak-through power. However, the effect n the other properties of the filter is negligible, and little diffi-ulty has been experienced in practice. Two of these units ere in service for some while, and apparently gave very little ouble. Such failures as occurred could not be attributed to printed circuit. Some anxiety, however, was felt about the fects of moisture condensing on the surface of the printed nel, but it appears that hermetic sealing of the unit as a whole vides adequate protection.

(8) CONCLUSIONS

A wide variety of microwave components have been described, of which are in strip-line form, and fabricated by printed-circuit techniques. These components have a performance which is generally comparable with waveguide; the r.f. losses tend to be slightly higher, but the operating bandwidth is usually larger. There is also a limitation in regard to power-handling capacity.

For broad-band non-resonant applications the open microstrip system has been found quite satisfactory. It has the merit of simplicity and is easier both to test and to manufacture. There is a small tendency to radiate, however, and microstrip receiver components should therefore not be exposed to large r.f. fields, such as might be produced by a badly screened local oscillator or a transmitter. These components must also be protected against the weather; with p.t.f.e. Fibreglass, however, the water absorption is quite low and prevention of excessive surface condensation is probably all that is required.

In the case of resonant circuits the high-Q triplate line is perhaps more satisfactory, although this does not exclude the possibility of certain types of filter circuits in microstrip. Triplate line is also rather more convenient where a totally enclosed packaged component with plug-in coaxial connections is required. The two ground planes, together with the side and end supports, form a natural enclosure, whereas in the case of microstrip a special box must be provided and certain precautions may be necessary to prevent resonances in this box.

It is too early to comment other than very generally on the economics of strip-line components. The cost of fabrication appears to be considerably less than for normal waveguides, whilst the cost of testing the components is likely to be very much the same. Generally speaking, it is safe to say that the economic advantage will be much larger than in the case of low-frequency printed-circuit applications, where the advantage is usually very marginal except when very large quantities and extensive mechanization are involved. For this reason the use of strip-transmission-line techniques is already proving economically worth while, even when the quantities involved are of the order of hundreds or even dozens. In certain applications, however, the economic advantage may be quite subordinate to certain other features, e.g. small size, bulk and weight, inherently large bandwidth, and a potentially high order of reproducibility.

(9) ACKNOWLEDGMENTS

The author would like to acknowledge the assistance of a number of colleagues, notably Mr. L. Lewin, who supervised much of the work, Mr. D. A. Williams, for microwave experiments and fabrication of specimens, Mr. J. Leno, for advice on printed-circuit techniques, Mr. W. E. Simpson, for preparation of special materials, and others less specifically associated with the work reported, notably Mr. J. Kemp. Acknowledgment is also due to the staff of the Federal Telecommunication Laboratories, New Jersey, who made available much valuable information concerning their own developments in this field. The paper is published by permission of Standard Telecommunication Laboratories Limited.

(10) BIBLIOGRAPHY

(1) GRIEG, D. D., and ENGELMANN, H. F.: 'Microstrip—A New Transmission Technique for the Kilomegacycle Range', *Proceedings of the Institute of Radio Engineers*, 1952, **40**, No. 12, p. 1644.

(2) ASSADOURIAN, F., and RIMAI, E.: 'Simplified Theory of Microstrip Transmission Systems', *ibid.*, p. 1651.

(3) KOSTRIZA, J. A.: 'Microstrip Components', *ibid.*, p. 1658.

(4) BARRETT, R. M.: 'Microwave Printed Circuit—A Historical Survey', *Transactions of the Institute of Radio Engineers*, March, 1955, MTT-3, p. 1.

(5) FROMM, W. E.: 'Characteristics and Some Applications of Strip-line Components', *ibid.*, p. 13.

(6) WILD, N. R.: 'Photoetched Microwave Transmission Lines', *ibid.*, p. 21.

(7) COHN, S. B.: 'Problems in Strip Transmission Lines', *ibid.*, p. 119.

(8) RINGENBACH, M. E., and COOPER, H. W.: 'Measurement of Attenuation and Phase Velocity of Various Laminate Materials at L-Band', *ibid.*, MTT-3, p. 87.

(9) DUKES, J. M. C.: 'An Investigation into some Fundamental Properties of Strip Transmission Lines with the Aid of an Electrolytic Tank', *Proceedings I.E.E.*, Paper No. 1991 R, May, 1956 (**103** B, p. 319).

(10) FOSTER, K.: 'Characteristic Impedance of Some Forms of Shielded Strip Transmission Line', A.C. Cossor Internal Memorandum.

(11) ARDITI, M.: 'Characteristic and Application of Microstrip for Microwave Wiring', *Transactions of the Institute of Radio Engineers*, March, 1955, MTT-3, p. 31.

(12) OLINER, A. A.: 'Theoretical Developments in Symmetrical Strip Transmission Line', *Proceedings of the Symposium on Modern Advances in Microwave Techniques*, Polytechnic Institute of Brooklyn, November, 1954, p. 379.

(13) FUBINI, E., FROMM, W., and KEEN, H.: 'New Technique for High-*Q* Strip Microwave Components', *Convention Record, Institute of Radio Engineers*, 1954, Part 8, p. 91.

(14) ATTWOOD, S. S.: 'Surface Wave Propagation over a Coated Plane Conductor', *Journal of Applied Physics*, 1951, **22**, p. 504.

(15) BRUNETTI, C., and CURTIS, R. W.: 'Printed Circuit Techniques', National Bureau of Standards, 1947, Circular 468.

(16) 'New Advances in Printed Circuits', National Bureau of Standards, Miscellaneous Publication 192, 1948.

(17) Proceedings of the Symposium of the Radio Electronics Television Manufacturers Association, New York. Held at University of Pennsylvania, January, 1955.

(18) BOWNESS, C.: 'Strip Transmission Lines', *Electronic Engineering*, January, 1956, **28**, No. 335, p. 2.

(19) SCIEGIENNY, J.: 'Strip above Ground Plane Transmission Systems', Quarterly Progress Report of the Research Laboratories of Electronics, M.I.T., Part x, January, 1953, p. 53.

(20) ALTSCHULER, H. M., and OLINER, A. A.: 'Microwave Measurements with a Lossy Variable Termination', *Proceedings I.E.E.*, Monograph No. 179 R, May, 1956 (**103** C, p. 392).

(21) DUKES, J. M. C.: 'Waveguides and Waveguide Junctions—Geometrical Analysis of their Properties', *Wireless Engineer*, March, 1955, **32**, p. 65.

(22) DUKES, J. M. C.: 'Transmission Line Termination—Measurement through a Mismatched Junction', *ibid.*, October, 1955, **32**, p. 266.

(23) DESCHAMPS, D. A.: 'Determination of Reflection Coefficients and Insertion Loss of a Waveguide Junction', *Journal of Applied Physics*, 1953, **24**, p. 1046.

(24) ZUBLIN, K. E.: 'Strip Type Components for 2000 Mc. Receiver Head End', *Transactions of the Institute of Radio Engineers*, March, 1955, MMT-3, p. 65.

(25) DUKES, J. M. C.: 'Broad-band slot coupled Microstrip Directional Couplers' (see page 147).

(26) SOMMERS, D. J.: 'Slot Array Employing Photoetched Tri-Plate Transmission Lines', *Transactions of the Institute of Radio Engineers*, March, 1955, MTT-3, p. 157.

(27) FUBINI, E., FROMM, W., and KEEN, H.: 'Microwave Applications of High-*Q* Strip Components', *Convention Record, Institute of Radio Engineers*, 1954, Part 8, p. 98.

(28) TORGOW, E. N., and GRIEMSMANN, J. E.: 'Miniature Strip Transmission Line for Microwave Applications', *Transactions of the Institute of Radio Engineers*, March, 1955, MTT-3, p. 57.

(29) WHITE, D. R., and BRADLEY, E. H.: 'Band-Pass Filters using Stripline Techniques', *ibid.*, p. 163.

(30) ALFORD, A.: 'Transmission Line Networks', *Electrical Communication*, January, 1939, **17**, No. 3, p. 301.

(31) FUBINI, E. G.: 'Stripline Radiators', *Transactions of the Institute of Radio Engineers*, March, 1955, MTT-3, p. 149.

(32) BITTNER, B. J.: 'Photo-etching Techniques for the Precision Fabrication of Waveguide Antenna Arrays', Symposium on Microwave Printed Circuits, Tufts College, October, 1954.

(33) TYMINSKI, W.: 'A Wide-Band Ring for U.H.F.', *Proceedings of the Institute of Radio Engineers*, 1953, **41**, No. 1, p. 81.

(34) GASCOYNE, D. N.: 'Reinforced Plastics—Behaviour in the Tropics', Report No. 470, Tropical Testing Establishment (Port Harcourt, Nigeria), Ministry of Supply, June, 1956.

[The discussion on the above paper will be found on p. 180.]

.1.372.821 : 621.372.542.2

The Institution of Electrical Engineers
Paper No. 2444 R
Nov. 1957

RE-ENTRANT TRANSMISSION-LINE FILTER USING PRINTED CONDUCTORS

By J. M. C. DUKES, M.A., Associate Member.

he paper was first received 21st June, and in revised form 5th September, 1957. It was published in November, 1957, and was read before the RADIO AND TELECOMMUNICATION SECTION *13th November, 1957.)*

SUMMARY

The re-entrant line filter section first described by Alford in 1939 employed. The theory of these filters is expanded, and a novel ocedure is described for the design of microwave low-pass filters ssessing a relatively high stop-band insertion loss over bandwidths 3 :1 or more. The filter may conveniently be realized in strip-nsmission-line form utilizing printed-circuit techniques. In this rm it appears to possess certain advantages over strip-line filters of re conventional design, notably the large operating bandwidth, d the accuracy with which the cut-off frequency can be computed. variety of these filters have been constructed both in microstrip and plate line, the latter form being more suitable for packaged units. very high degree of electrical reproducibility comparable with the st waveguide practice has been observed, with the advantage of duced manufacturing costs.

LIST OF PRINCIPAL SYMBOLS

Z_0 = System operating impedance, e.g. load or generator impedance, generally taken as unity.
Z_1, Z_2 = Characteristic impedance of transmission lines 1 and 2.
Z = Characteristic impedance of transmission lines 1 and 2 for $Z_1 = Z_2$.
Z_{sc}, Z_{oc} = Short-circuit and open-circuit impedances of complete filter section.
Z_{sch}, Z_{och} = Short-circuit and open-circuit impedances of a bisected section.
Z_i = Image impedance of filter section.
Z_{i0} = Zero-frequency image impedance of filter section.
Z_{ik} = Image impedance normalized to give $Z_{i0} = Z_0 = 1$.
γ = Propagation coefficient of whole filter section $(= \alpha + j\beta)$.
α = Attenuation coefficient of whole filter section.
β = Phase-change coefficient of whole filter section.
θ_1, θ_2 = Lengths of transmission lines 1 and 2.
f_c = Cut-off frequency.

(1) INTRODUCTION

The filter to be described was developed as part of an investigation into the properties of strip transmission lines,[1] the specific requirement being a low-pass filter with a sharp cut-off, an operating range of several octaves and a high degree of reproducibility. Since moderately large quantities might be required, the adoption of a printed-circuit technique had obvious attractions.[2]

The conventional low-pass filter circuit [Fig. 1(*a*)] was considered, with C and L replaced by short lengths of transmission line [Fig. 1(*b*)], but two major problems were encountered. First, in a planar system with uniform ground spacing it is difficult to obtain sufficiently large values of C and small values of L; the transmission lines accordingly tend to be rather long, and very critical proportioning of the circuit elements is required to prevent the first spurious pass-band occurring below $3f_c$. This difficulty is accentuated when m-derived terminations are to be added, for these tend to be even more frequency-sensitive than the prototype constant-k sections. The second problem arises from the important discontinuity effect at the junction between the series and shunt lines. For example, should the length of the series arm be measured between the centre-lines of the shunt strip conductors, or between the adjacent edges? Attempts to determine the cut-off frequency more precisely, using data given by Oliner,[3] have not proved entirely successful, and no alternative to trial and error has been found. Finally, the presence of the open end to the shunt arm is objectionable when microstrip is used, because it has a tendency to radiate.

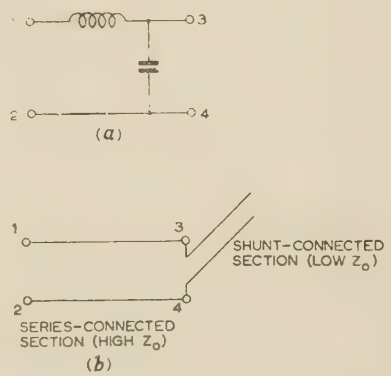


Fig. 1.—Prototype constant-k section.
(*a*) Low-frequency lumped circuit.
(*b*) High-frequency approximation using transmission lines.

These difficulties suggested that the filter should be approached entirely in terms of long-line theory, and of the various alternatives, the re-entrant line section (Fig. 2) appeared to have particularly attractive possibilities, in view of its similarity to conventional hybrid-ring structures. In experiments with microstrip hybrid rings it had already been found[4] that the effective electrical length of the ring is very closely equal to that calculated on the basis of the arithmetic mean circumference, even when the strip-conductor width is a considerable fraction of the ring diameter. It had also been found that both the junction effect at the connection to the ring and the tendency to radiate were small. An additional advantage appeared to be the absence of open-circuits and large changes in strip-conductor width, since the most favourable condition of operation is given by $Z_1 = Z_2$.

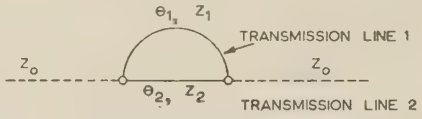


Fig. 2.—Single-section re-entrant line filter.

The re-entrant line circuit was first described by Alford[5] in connection with the matching of open-wire feeders, but because of the restricted character of his problem, Alford's analysis is only of limited use in the present application; the problem has been thus approached afresh, and a new technique has been

Mr. Dukes is with Standard Telecommunication Laboratories Limited.

developed for the design of multi-section filters giving results equivalent to those obtained by a conventional m-derivation process.

(2) PRINCIPLES OF OPERATION

(2.1) The Single Re-Entrant Section

The elementary filter section shown in Fig. 2 is analysed in Section 6.1, where it is shown that, if a high stop-band insertion loss is required over a large bandwidth, the most favourable operating condition is given by $Z_1 = Z_2 = Z$; the expressions for the image impedance and the propagation coefficient derived there may then be simplified to

$$Z_i = \tfrac{1}{2}Z\left[\frac{\sin\theta_1 \sin\theta_2}{\sin^2 \tfrac{1}{2}(\theta_1 + \theta_2)}\right]^{1/2} \quad . \quad . \quad . \quad . \quad (1)$$

and

$$\gamma = \operatorname{arc\,tanh}\left[-\frac{\sin\theta_1 \sin\theta_2}{\cos^2 \tfrac{1}{2}(\theta_1 + \theta_2)}\right]^{1/2} \quad . \quad . \quad . \quad (2a)$$

or

$$\gamma = 2 \operatorname{arc\,tanh}(-\tan\tfrac{1}{2}\theta_1 \tan\tfrac{1}{2}\theta_2)^{1/2} \quad . \quad . \quad (2b)$$

If θ_1 is the longer line, the principal cut-off frequency is given by the condition $\theta_1 = \pi$; the other cut-off points, together with the frequency for which $\alpha = \infty$, can be determined by inspection of eqns. (1) and (2), the latter being given by $(\theta_1 - \theta_2) = (2n + 1)\pi$. These data are presented in graphical form in Fig. 3,

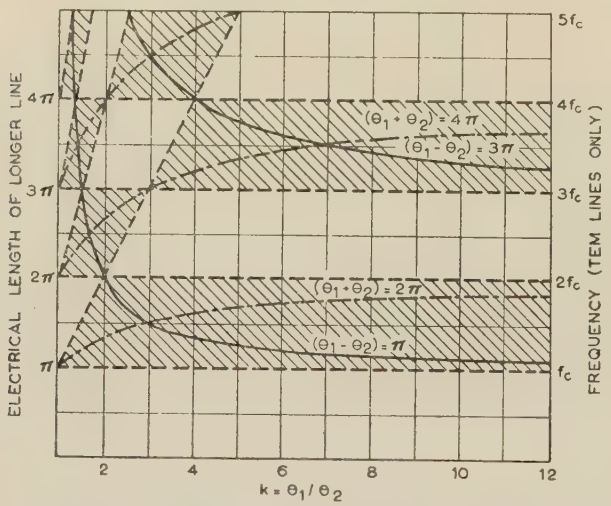


Fig. 3.—Location of stop bands and pass bands.

——— Locus of $\alpha = \infty$.
— — — — Locus of $Z_i = 0$.
—·—·—·— Locus of $Z_i = \infty$.
Shaded areas denote stop bands.

which also shows the zeros and poles of Z_i. From eqn. (3) the zeros occur either when $\theta_1 = n\pi$ or when $\theta_2 = n\pi$, and the poles occur when $(\theta_1 + \theta_2) = 2n\pi$.

It will be noted that, as with all transmission-line filters, the stop bands are periodic with frequency; furthermore, by a suitable choice of parameters, the width of the stop bands or pass bands can be adjusted as required. Hence the filter is of the most general kind and may be used for low-pass, high-pass, band-pass or band-stop operation.

Now it is important to note the manner in which Z_i varies with θ_1/θ_2. Let $\theta_1/\theta_2 = k$, a constant; then as the frequency tends to zero, Z_i tends to

$$Z_{i0} = Zk^{1/2}(1 + k)^{-1} \quad . \quad . \quad . \quad . \quad . \quad (3)$$

It will be observed that the image impedance falls with increasing disparity in the line length. In general, it is desirable to se Z_{i0} equal to the load and generator impedance, Z_0, i.e. so that

$$Z = Z_0(1 + k)k^{-1/2} \quad . \quad . \quad . \quad . \quad . \quad (4$$

At the same time it is convenient to normalize the impedanc level by setting $Z_0 = 1$, the normalized image impedance, Z_{ik} being given by

$$Z_{ik} = \tfrac{1}{2}(1 + k)k^{-1/2}\frac{[\sin\theta_1 \sin(\theta_1/k)]^{1/2}}{\sin\tfrac{1}{2}(1 + k^{-1})\theta_1} \quad . \quad . \quad (5$$

This relationship is plotted in Fig. 4. Comparison of these curves with those in Fig. 3 shows that the closer the peak of infinite attenuation to the cut-off frequency the worse is the match just below the cut-off frequency.

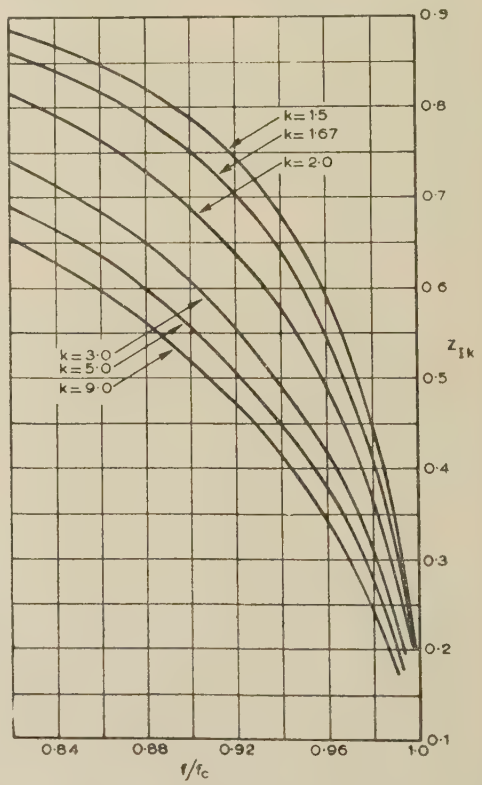


Fig. 4.—Variation of image impedance just below cut-off.
All sections normalized to have same zero-frequency image impedance.

(2.2) Filters with Several Identical Sections in Cascade

In the simplest case, where the specification is not too strict it may be adequate to employ a number of identical sections in cascade without the aid of special terminating sections. In order to achieve the best possible pass-band characteristic, a low value of k should be used, although—as can be seen from Fig. 3—a value of k less than 2 will result in a reduced bandwidth for the first stop band. In this connection it will be noted that the value $k = 2$ appears particularly attractive, since the stop band extends from f_c to $3f_c$. A design based on this value is described in Section 3.

(2.3) Filters with Transforming End Sections

For reasons discussed in Section 7.2, the conventional m-derivation technique is not applicable to this type of filter. An alternative procedure has been outlined by Cohen[6] for the conventional prototype low-pass section whereby the end section

ve a cut-off frequency approximately 1·3 times that of the ntral sections. This is based on the fact that at $0{\cdot}707f_c$ the ase shift of a prototype section is 90° and its impedance is $.707Z_0$. Thus, at the frequency just below cut-off of the ntre sections, where the impedance is $\frac{1}{2}Z_0$, the end section erates as a quarter-wave matching section. The same tech-que can be used here, but because Z_i and γ are functions of k well as frequency, the precise numerical factor would depend the value of k employed. This additional freedom has ggested a variation of this technique which has the advantage at the cut-off frequency is the same for both the end and the ntre sections. The centre sections are chosen to have a fairly gh value of k, as a result of which Z_i is low in the cut-off region. compensate for this, the two end sections are each given a w value of k such that at some frequency below the cut-off quency the image impedance of the end section is the geometric an of the generator/load impedance and the centre-section age impedance, the numerical values being determined from g. 5. The two values of k are selected by trial and error, so

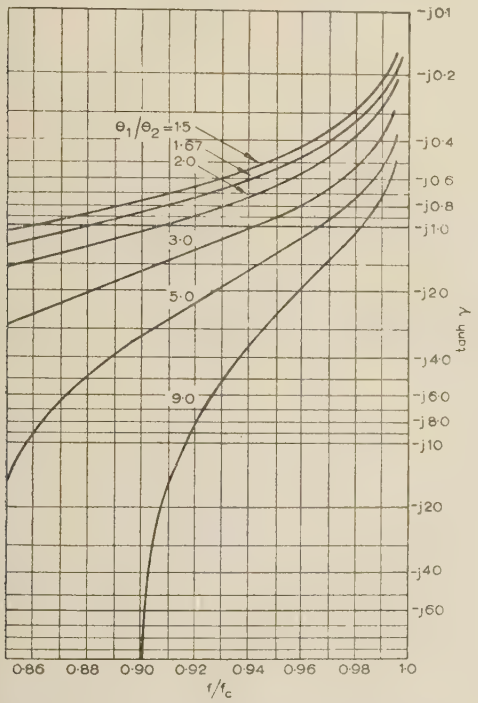


Fig. 5.—Variation of tanh γ just below cut-off.

iat, at the frequency at which the geometric-mean relationship maintained, the phase shift of the end section is 90°, i.e. nh $\gamma = \infty$. The precise matching conditions are maintained one frequency only, but because the image-impedance functions r the two values of k vary with frequency in the same general anner, and because the tanh γ function for the low value of k rather flat, an improved match is obtained over quite a sub-antial bandwidth.

(2.4) The Tapered Re-Entrant Line Filter

The technique described briefly in the latter part of Section 3 may be improved by providing a gradual variation in k roughout the length of the filter, as illustrated schematically Fig. 6. The inner sections, which are given high values of k, ovide a sharp cut-off characteristic, whilst the low-k outer ctions ensure a good match to the external line impedance at equencies just below the cut-off. In order to obtain a good

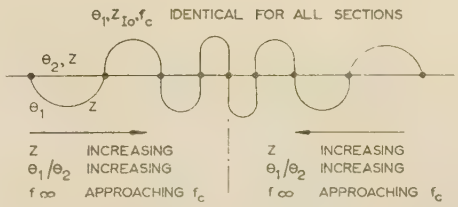


Fig. 6.—Schematic of tapered re-entrant line filter.

internal match at frequencies well below the cut-off, the zero-frequency image impedances are made identical for all sections, with the result that the transmission-line impedance tapers from a low value at the outer sections to a high value at the centre.

From Fig. 4 it will be seen that at frequencies just below cut-off there is a gradual fall in normalized image impedance as the centre of the filter is approached. The filter as a whole may therefore be regarded as a form of tapered transmission line, and may to some extent be analysed accordingly. The principal difference is that in the pass band the impedance gradient (i.e. the rate of change of impedance with respect to equivalent electrical length) increases with frequency.

The advantage of regarding the tapered re-entrant line filter in this way is that certain criteria already exist for ordinary tapered lines;[7] these may be used to determine qualitatively the number of transforming sections required and their impedance ratio in order that the v.s.w.r. shall not exceed a certain specified figure at some given frequency just below the cut-off.

So far as the stop-band performance is concerned, it would seem desirable to select values of k which will produce a fairly even spread of frequencies of infinite attenuation throughout the range of interest. There is no fundamental conflict between this requirement and that previously expressed in regard to the pass-band requirements, but it should be noted that a linear impedance taper will result in a rather severe crowding of the frequencies of infinite attenuation towards the lower edge of the principal stop-band. A more even spread would be produced by an exponential taper.

Since the individual sections are not matched on an image-impedance basis, the detailed analysis of the complete filter is extremely difficult. Matrix methods may be used,[7] but the expressions finally obtained are so complicated as to give virtually no clue to overall performance without resort to laborious numerical evaluation. Numerical resolution has, in fact, been undertaken only in order to check certain specific regions of interest.

The inability to match all sections on an image-impedance basis introduces another difficulty, in that, under certain circumstances (see Section 7.4) the mismatch between adjacent sections can cause spurious responses in the stop band. Methods of overcoming this difficulty are discussed in Section 3.2.

(3) EXPERIMENTAL WORK

(3.1) Simple Filter

An experimental design based on six identical sections and fabricated entirely in microstrip (open-type line-above-ground) is illustrated in Figs. 7 and 8, each section having $k = 2$. In order to improve the match at the upper end of the pass band, the operating impedance level has been increased somewhat above that appropriate to the condition $Z_{i0} = Z_0$. The rate of cut-off of this design compares quite favourably with the more elaborate design described later. On the other hand, the pass-band insertion loss is somewhat high, and the stop band, extending as it should from 3 to 9 Gc/s, is punctuated by a number of very narrow spurious responses. These responses are not unique to this

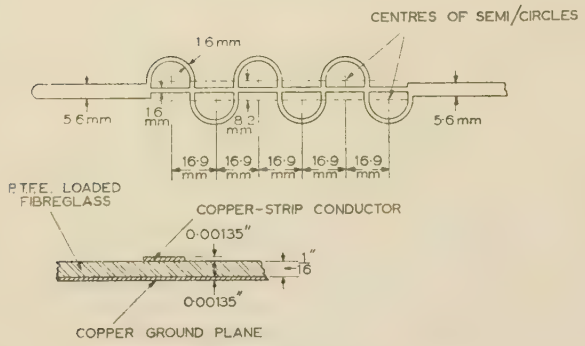


Fig. 7.—Six-section low-pass filter.

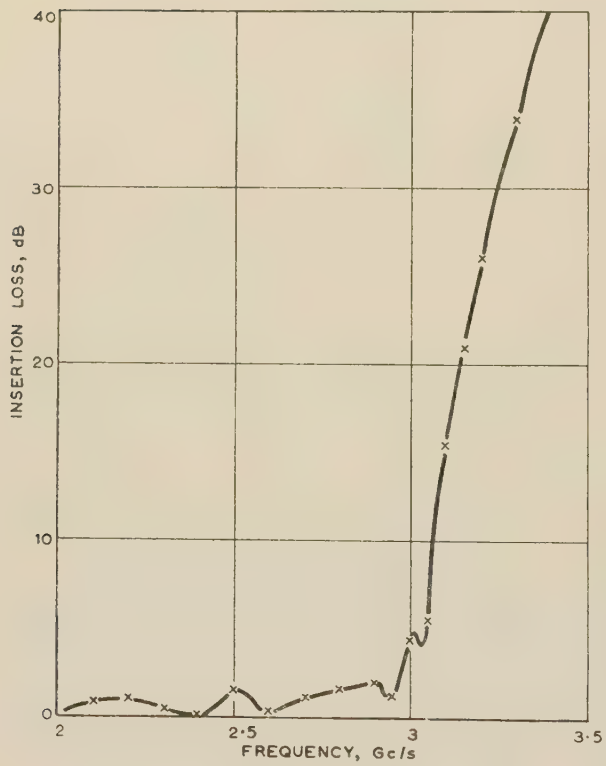


Fig. 8.—Response of 6-section low-pass filter.

design, however, and they are discussed in greater detail in Section 3.2.

No attempt was made to cure the spurious response, since the filter suffered from a rather more serious limitation of a practical character: owing to the open nature of the line, the performance of the filter in the cut-off region was seriously affected by metallic objects in its vicinity; the sensitivity depended on the orientation of the object with respect to the axis of the line, and was most severe in a direction vertically above the line, where a clearance of at least 10 in was required. At frequencies remote from cut-off (either stop-band or pass-band) the sensitivity was very much reduced. It is reasonable to assume that in view of the sharp cut-off a small change in energy stored in the far field was sufficient to produce a noticeable detuning effect.

Filters with simple end transforming sections as described in Section 2.3 have also been constructed. Worth-while improvements were obtained, and the method appears to have considerable application where a short filter is required.

(3.2) Tapered Re-Entrant Line Filter

A considerable amount of work was undertaken on this ty of filter, mostly devoted to eliminating spurious stop-ba responses and adjusting the performance by the elimination addition of one or two sections. Basically there was no alter tion to the initial design, and it is noteworthy that on the fir trial the measured cut-off frequency was within 2% of t calculated value. The high-Q-factor triplate[2] form of co struction was used (see Fig. 9), thus avoiding the difficulti

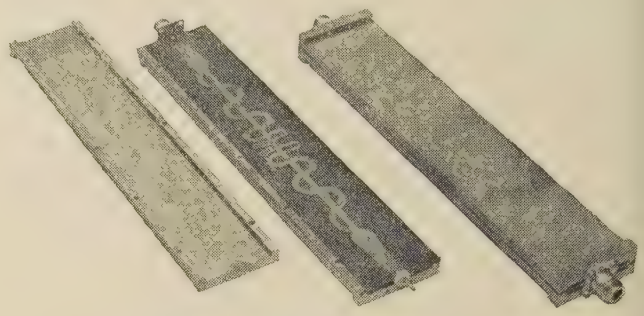


Fig. 9.—Tapered re-entrant line filter.
The high-Q-factor triplate form of construction is used.

mentioned in Section 3.1, the Fibreglass sheet being supporte between two U-shaped channels which ensure that the groun plane spacing follows any variation in the sheet thickness. Und these conditions the impedance variation is less than if t ground-plane spacing were held constant. To ensure adequa contact between the two ground planes, copper-foil strip is folde around each edge of the Fibreglass sheet. The basic parameter of this design are summarized in Table 1.

Table 1

SUMMARIZED DESIGN DATA FOR TAPERED RE-ENTRANT LIN FILTER

Nominal cut-off frequency	4 100 Mc/s
Operating range	2 500–11 500 Mc/s
Ground-plane spacing	5/16 in
Thickness of dielectric sheet ..	1/16 in
Dielectric-sheet material	P.T.F.E.-impregnated Fibreglas
Strip conductor	0·001 35 in copper foil
Ground planes	Brass
Total length of filter, including tapered end-sections	12½ in

Section Nos.	1, 15	2, 14	3, 13	4, 12	5, 11	6, 10	7, 9	8
θ_1/θ_2 ..	1·33	1·5	1·67	2·0	3·0	5·0	9·0	17·0
f/f_c	4	3	2·5	2·0	1·5	1·25	1·125	1·06
Z_0	60·6	61·2	62·1	63·6	69·0	81·8	99·9	132·

It will be noted that the design is based on a load impedanc of 30 ohms, which was chosen to reduce the impedance of th centre sections and hence maintain a reasonable strip-conducto width. Even so, the width of centre-section conductor is onl 0·017 in, which is about the limit of reliable photo-etching. A each end of the filter a tapered line is therefore provided to mat the filter to the external coaxial-line impedance of 50 ohm the length of these tapers being approximately $\lambda/2$ at the mi pass-band frequency of 3·3 Gc/s.

To test the reproducibility of these filters, four complete uni were manufactured; the ground-plane U-channels were manu factured to a tolerance of approximately 0·001 in, but the coaxi fittings were picked at random from stock. The four pieces Fibreglass were cut from the same sheet, but otherwise n selection process was involved. To ensure accurate registratio of the front and reverse prints, glass plates were used in co

Table 2

REPRODUCIBILITY TEST ON FOUR FILTERS

requency, Gc/s	2·47	2·76	3·07	3·45	3·66	3·86	3·98	4·04	4·135	4·14	4·24
oread, dB	0·29	0·43	0·24	0·36	0	0·20	0·47	0·64	0·37	2·97	2·37

nction with a special arrangement of dowel pins. The prints ere then X-rayed to determine the accuracy of registration, hich was found to be within 0·005 in. With the method of inting employed, the overall sizes of the circuits must, of urse, be the same, but small discrepancies in the width of the nductor can arise from undercutting of the resist by the etching lution. With careful workmanship the variation should not ceed a small fraction of 0·001 in, except perhaps with very in conductors, e.g. 0·005 in, where the resist may peel off mpletely, owing to insufficient care in handling.

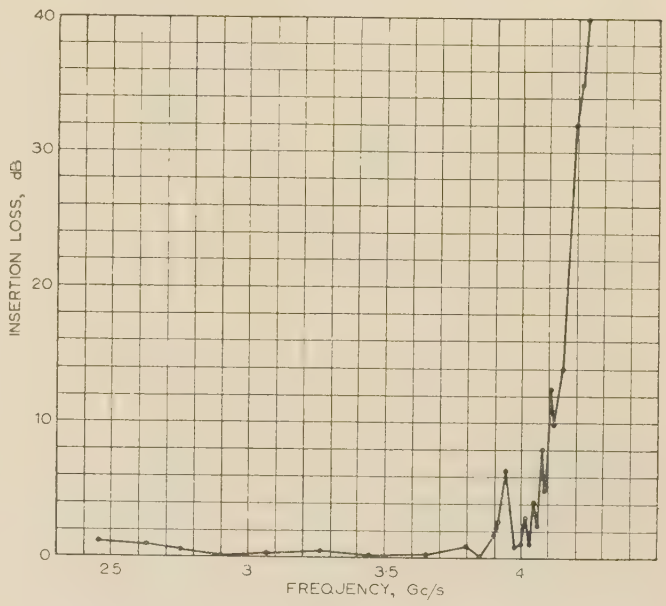
Owing to faulty photo-etching (fortunately a relatively rare currence) two of the four filters were imperfect, the foils being tted with small holes and scratches; in one case the imperfections were repaired by soldering, and in the other by silver ating. The final results of the reproducibility test are summarized in Table 2, and it will be observed that the four filters are nsistent within $\frac{1}{2}$dB right up to the cut-off frequency. It is lieved that, with a carefully controlled etching process and reful mechanical inspection, the same order of reproducibility ould be possible under quantity production.

As already mentioned, much experimental work was devoted the elimination of spurious responses in the stop band, which curred principally at about 7 and 10·6 Gc/s and were stematically reproducible. A number of possible causes ere considered and investigated, including spurious coupling etween the input and output by the H_{0n} or other modes,[2] apacitative or inductive coupling between adjacent conductors nd possible spurious effects arising, for example, from the vision of the centre conductor into separate upper and lower nductors. Since the results of these tests were largely negative, was finally concluded that the spurious responses were an herent weakness of the design technique.

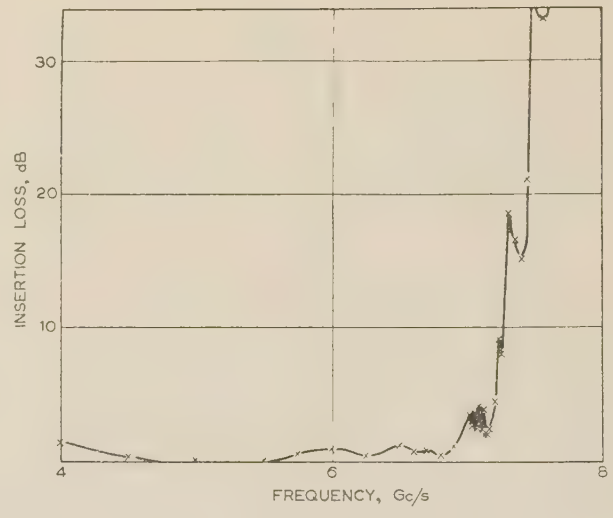
The latter possibility was investigated fairly carefully, first by direct numerical computation and secondly via the restricted nalysis given in Section 6.3. Ten points in certain critical parts f the spectrum were computed numerically, but did not reveal ny spurious response. This is not of great significance, since e measured responses are very narrow (e.g. 50 Mc/s between 40 dB points) and the effort required to investigate the entire op-band of 8 Gc/s in this degree of detail would be quite rohibitive (unless undertaken on a computer). By contrast, the stricted analysis given in Section 6.3 suggests that at least a few sponses are probable, and some criteria are there deduced as where they are likely to occur in the spectrum.

These conclusions would seem to be supported by the result f experiments in which the conductors and ground planes are lated with metals having a relatively poor conductivity; e.g. ith tin the transmission at the peak of the spurious response ay be reduced by as much as 15 dB or more. This technique as virtually no effect on the pass-band response, and is therefore ery useful, although by itself it does not appear to be sufficient. was therefore found necessary to add a simple choke-type ffer at each end. These are integral with the tapered matching ctions, thus maintaining the same overall length of the filter. he measured results are given in Fig. 10(*a*).

In addition to the filter with a cut off at 4·1 Gc/s, a scaled rsion with a cut-off at 7 Gc/s was also constructed, with the sults shown in Fig. 10(*b*). Owing to the restricted width of the minal stop-band (7–11·5 Gc/s) no difficulty was experienced th spurious responses, and the addition of choke sections was sidered unnecessary.



(*a*)



(*b*)

Fig. 10.—Measured responses of filters.
(*a*) 4 Gc/s.
(*b*) 7 Gc/s.
From 4 to 11 Gc/s the loss is greater than 20 dB.

(4) CONCLUSIONS

The experiments have shown that satisfactory microwave filters can be constructed by printed-circuit methods. The low-pass filters described have electrical properties equal to those of similar coaxial or waveguide components, with the advantage of lower production costs and an inherently high order of reproducibility. A new type of filter circuit has been devised which, although unsuitable for other forms of construction, is particularly well adapted to the strip-transmission-line technique.

The success achieved suggests that other types of filter, e.g. high-pass, band-pass, etc., should be possible, but it is likely that the best results, as in this case, will be achieved by selecting resonant structures which are in some degree particular to strip lines rather than by imitating existing waveguide or coaxial forms.

(5) ACKNOWLEDGMENTS

The author would like to thank Mr. L. Lewin for a number of helpful discussions, and Messrs. D. A. Williams and F. R. Sagnard, who carried out most of the experimental work. The paper is published by permission of Standard Telecommunication Laboratories.

(6) REFERENCES

(1) DUKES, J. M. C.: 'The Application of Printed-Circuit Techniques to the Design of Microwave Components' (see page 155).
(2) 'Symposium on Microwave Strip Circuits', *Transactions of the Institute of Radio Engineers*, March, 1955, MTT-3, No. 2.
(3) OLINER, A. A.: 'Equivalent Circuits for Discontinuities in Balanced Strip Transmission Lines', *ibid.*, p. 134.
(4) *Ibid.*, p. 31.
(5) ALFORD, A.: 'Transmission-Line Networks', *Electrical Communication*, 1939, **17**, p. 301.
(6) 'Very High Frequency Techniques' (McGraw-Hill, New York, 1947), p. 666.
(7) RAGAN, G. L.: 'Microwave Transmission Circuits', M.I.T. Radiation Laboratory, Laboratory Series No. 9 (McGraw-Hill, New York, 1948).

(7) APPENDICES

(7.1) Basic Analysis of Filter

Despite the simplicity of the circuit, a direct attack on the filter leads to very lengthy algebraic manipulations, except in the special case when $Z_1 = Z_2$, and it is simpler to proceed by application of Bartlett's bisection theorem,[7] as follows. For the short-circuit termination the half-circuit input impedance is given by

$$Z_{sch}^{-1} = (jZ_1 \tan \tfrac{1}{2}\theta_1)^{-1} + (jZ_2 \tan \tfrac{1}{2}\theta_2)^{-1} \quad . \quad . \quad (6)$$

and for the half-circuit open-circuit impedance by

$$Z_{och}^{-1} = (-jZ_1 \cot \tfrac{1}{2}\theta_1)^{-1} + (-jZ_2 \cot \tfrac{1}{2}\theta_2)^{-1} \quad . \quad (7)$$

By Bartlett's bisection theorem the image impedance for the complete section is given by

$$Z_i = (Z_{sch}Z_{och})^{1/2} \quad . \quad . \quad . \quad . \quad . \quad . \quad (8)$$

$$= Z_1Z_2[Z_1 \tan \tfrac{1}{2}\theta_1 + Z_2 \tan \tfrac{1}{2}\theta_2)(Z_1 \cot \tfrac{1}{2}\theta_1 + Z_2 \cot \tfrac{1}{2}\theta_2]^{-1/2} \quad . \quad . \quad . \quad . \quad . \quad . \quad (9)$$

Likewise, the propagation coefficient for the network as a whole is given by

$$\gamma = 2 \text{ arc tanh } (Z_{sch}/Z_{och})^{1/2} \quad . \quad . \quad . \quad . \quad (10)$$

$$= 2 \text{ arc tanh } \left(-\frac{Z_1 \tan \tfrac{1}{2}\theta_2 + Z_2 \tan \tfrac{1}{2}\theta_1}{Z_1 \cot \tfrac{1}{2}\theta_2 + Z_2 \cot \tfrac{1}{2}\theta_1}\right)^{-1/2} \quad . \quad (11)$$

Assuming the transmission lines to be free from dissipation, pass bands occur when Z_i is real and γ imaginary, and the stop bands occur when γ is real and Z_i is imaginary.

The frequencies at which the attenuation of the section is infinite are given by the condition $\tanh \tfrac{1}{2}\gamma = 1$.

The principal effect of varying Z_1/Z_2 may be illustrated in the following manner. From eqn. (9) it can be seen that whe $Z_1 = Z_2$ the condition $(\theta_1 + \theta_2) = 2\pi$ is associated with a pol of the image-impedance function Z_i; Z_{sch} and Z_{och} likewise bot pass through a pole simultaneously at this frequency. Howeve if $Z_1 \neq Z_2$ the poles of Z_{sch} and Z_{och} are displaced, but i opposite directions, which introduces a narrow pass-ban centred on the frequency corresponding to $(\theta_1 + \theta_2) = 2\pi$. Generally, the presence of these narrow responses restricts th application of the filter and it is therefore desirable in mos cases to make $Z_1 = Z_2$.

(7.2) Limitations of the *m*-Derivation Procedure

As is well known, the image-impedance and propagation coefficients are related to the short-circuit and open-circui impedances by the relationships

$$Z_i = (Z_{sc}Z_{oc})^{1/2}$$

$$\tanh \gamma = (Z_{sc}/Z_{oc})^{1/2}$$

For lumped-circuit filters Z_{sc} and Z_{oc} are polynomials in ω and it is always possible to multiply all terms in Z_{sc} by a constan m and at the same time to divide all terms in Z_{sc} by the sam constant; thus

$$Z_i = (Z_{sc}Z_{oc})^{1/2} \text{ as before}$$

but

$$\tanh \gamma = m(Z_{sc}/Z_{oc})^{1/2}$$

This is the process of m-derivation,[7] the features and advantage of which are well known.

For transmission-line filters such as the present one, Z_{sc} and Z_{oc} must be expressed as polynomials in $\tan a\omega$ and $\cot a\omega$ where a is an appropriate constant. Strictly speaking, the m-derivation process is no longer possible, but over a limited band of frequencies it is often possible to make the approximations $\tan a\omega = a\omega$ and $\cot a\omega = a\omega^{-1}$, which, in effect, means that the transmission lines must be relatively short compared with a wavelength—which is evidently not generally true with the present filter, whose properties depend essentially on long-line resonance phenomena. Some further conclusions regarding this question are mentioned at the end of Section 7.3.

(7.3) Cascading Dissimilar Filter Sections

Consider two filter sections in cascade, both operating in a stop band; let the image impedances (both imaginary) be jX_1 and jX_2, and the hyperbolic tangents of the propagation coefficients (both real) μ_1 and μ_2, noting that only X can assume negative values. By the application of standard network theory it is a straightforward matter to show that the image impedance of the two sections in cascade is given by

$$Z_i = X_2\left[-\frac{X_1X_2\mu_2(1+\mu_1^2) + \mu_1(X_1^2 + \mu_2^2X_2^2)}{X_1X_2\mu_2(1+\mu_1^2) + \mu_1(X_2^2 + \mu_2^2X_1^2)}\right]^{1/2}$$

The condition for a spurious pass-band to occur is that Z_i shall be a real positive number. It is clear that this can occur only if X_1 and X_2 are of opposite sign, and only in the limited range where either the numerator or denominator, but not both, is negative.

To simplify the analysis divide through by $X_1X_2\mu_1\mu_2$ and let $X_1/X_2 = a$.

Then

$$Z_i = X_2\left(-\frac{\mu_1 + \dfrac{1}{\mu_1} + \dfrac{a}{\mu_2} + \dfrac{\mu_2}{a}}{\mu_1 + \dfrac{1}{\mu_1} + a\mu_2 + \dfrac{1}{a\mu_2}}\right)^{1/2}$$

ence the conditions for a spurious response may be expressed rmally as

(1) a must be negative.

(2)
$$\left|\frac{a}{\mu_2}+\frac{\mu_2}{a}\right| > \left(\mu_1+\frac{1}{\mu_1}\right) > \left|a\mu_2+\frac{1}{a\mu_2}\right|$$

$$\left|\frac{a}{\mu_2}+\frac{\mu_2}{a}\right| < \left(\mu_1+\frac{1}{\mu_1}\right) < \left|a\mu_2+\frac{1}{a\mu_2}\right|$$

Now the minimum value that any of the three expressions can sume is 2, and this occurs when any one of the terms a/μ_2, 2, or $a\mu_2$ is equal to unity. Hence it follows that no spurious sponse can occur if $\mu_1 = 1$, i.e. if the attenuation of section 1 infinite. Furthermore, if $\mu_2 = 1$, the first and last expressions e identical, and again no spurious response is possible. In th cases this is true whatever the value of a. Other conditions r absolute absence of spurious response are $\mu_1 = 0$, $\mu_2 = 0$, both, and $a = -1$.

Now from Foster's reactance theorem we know that the nctions $X_1(f)$ and $X_2(f)$ must both possess positive slope. o similar conclusion can be drawn in regard to $a(f)$, which ay accordingly have either positive or negative slope. There a restriction, however, that the poles and zeros shall alternate. If these considerations are applied to two adjacent sections

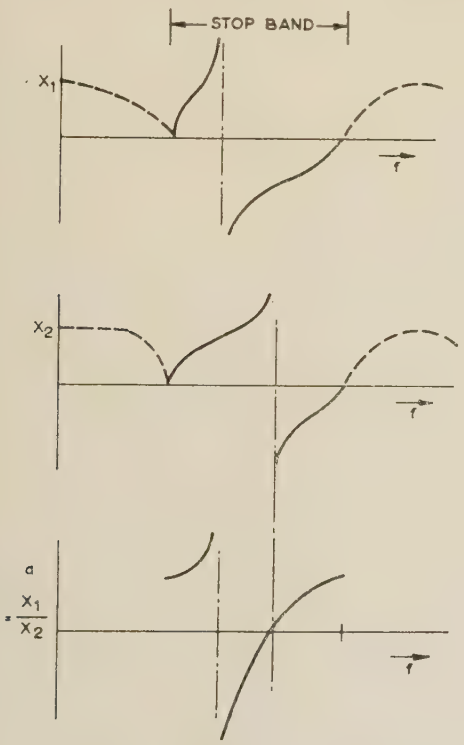


Fig. 11.—Image reactance curves.

the tapered re-entrant line filter, a diagram like Fig. 11 is tained, from which it will be seen that a passes through all ues from $-\infty$ to $+\infty$ and hence at least one spurious pass-nd must occur. Now in the region where a is negative, μ_1

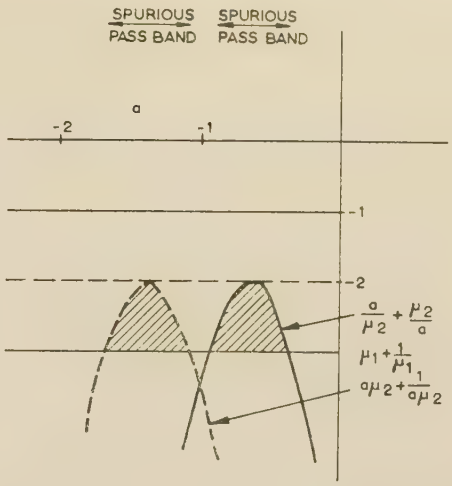


Fig. 12.—Appearance of spurious pass-bands.

and μ_2 are, in general, both close to unity and quasi-constant. Hence, if we plot the functions $|a/\mu_2 + \mu_2/a|$, $(\mu_1 + 1/\mu_1)$ and $|a\mu_2 + 1/(a\mu_2)|$ over the range where the spurious pass-band is likely to occur, a diagram like Fig. 12 will be obtained, which shows that, in general, two spurious pass-bands will occur, although, if $(\mu_1 + 1/\mu_1)$ is small enough, these two pass-bands may close up into one.

Three further general conclusions may be drawn. First, the two spurious pass-bands will be situated at very approximately the mid-frequency between the poles of the two image-impedance functions. Secondly, if the attenuation of the two sections is high, the spurious pass-bands will be very narrow. Thirdly, if the inherent Q-factor of the filter is lower than that deduced from the bandwidth of the spurious responses calculated for the loss-free case, the spurious responses will be smoothed out until, for practical purposes, they virtually disappear. This is in good agreement with experimental results (Section 3.3).

Formal application of this technique to the multi-section filter is difficult. From a general point of view it is clear that the number of spurious responses increases rather rapidly with the number of sections, but that the bandwidth of the responses is decreased. In the practical instance these responses will not prove troublesome, provided that the peaks of infinite attenuation are distributed fairly evenly throughout the stop band and that the inherent Q-factor of the filter is not too high.

This reasoning illustrates a general criterion applicable to all filters, namely that any attempt to eliminate a spurious response by the addition of a second filter is very likely to introduce a second spurious response. By a suitable choice of impedance function for the second filter it may be possible to ensure that the new spurious response is outside the frequency range of interest. Alternatively, if the Q-factor of either filter is sufficiently low, the response may be flattened to a degree that it is no longer important. Thus an m-derived terminating section will frequently produce a spurious response at a frequency not too remote from the cut-off frequency because of the impracticability at microwave frequencies of matching it to the prototype section over an indefinitely wide frequency band (see Section 6.2).

[The discussion on the above paper will be found overleaf.]

DISCUSSION BEFORE THE RADIO AND TELECOMMUNICATION SECTION, 13TH NOVEMBER, 1957

Mr. L. Lewin: One of the major shortcomings of microstrip arises in measurement. Owing to the inadequacy of the theory in relation to the field patterns it is impossible to assess—as it is for a waveguide—the effects of a sampling probe close to the strip. This difficulty led to the Deschamps technique of measurement through a junction; but although there is a neat geometrical method for interpreting the results, it can be a very tedious and time-consuming technique when many measurements must be made over a wide band. It is also unsuitable for display purposes. We have therefore tended to abandon it in practice and to match the junction as well as possible. We measure in waveguide and expect errors of the order of 0·1 in the reflection coefficient, and eventually finish with components which appear to have a v.s.w.r. better than 0·7 over a broad band.

Recent improvements have involved a sliding load with a raised tip giving a reflection coefficient of better than 1%, and junctions matched to 0·95 over limited wave bands using tapered ridge waveguides, leading to consistent measurements of 0·9 v.s.w.r. in strip-line. Although this figure might be capable of some further improvement, a current limitation is the commercial coaxial cable connectors, some of which have v.s.w.r.'s as poor as 0·7 over parts of the broad bands in which we are interested.

When a strip-line was first mooted it was widely believed that it would cheapen microwave assemblies enormously. It was only necessary to prepare a master diagram, photograph and print it on Fibreglass board and one had as many components as necessary, like a page of stamps. It was not as simple as that, partly because one does not normally want isolated pieces. A big advantage is that one can make what is essentially an integrated assembly with all the pieces printed together. There are no flanges or joints. In such circumstances the assembly as a whole is printed. So, although it is pretty to see a board with a large number of hybrid rings on it, such a method is of rather limited application.

For more detailed aspects one finds that it is not merely a question of printing; for example, coaxial or waveguide feeds must be inserted. The author gives an excellent example in his back-to-back directional coupler, which represents a significant advance in microwave technique as a whole rather than specifically strip-line. It is, however, no longer a straightforward printed circuit—several pieces must be made separately and soldered together. Coaxial connectors must be coupled at the ends—by which time one has gone through many of the motions necessary in providing an ordinary waveguide coupler and come a long way from the original idea of a simple print.

A similar situation occurs in attenuators when a slit is cut in the ground plane and an iron-dust load is cast in a sliding section. There must be some precisely made rails and slides for the load and controls for regulating its position, etc. By the time this has been done it is probably as expensive and cumbersome as a waveguide attenuator, of which there are many fairly simple designs. All this should not be taken as being against strip-line: it is merely that one should not run away entirely with the idea that it is printing and nothing more.

With regard to fields of application, an excellent example is the light-weight compact propagation testing set, compactness and light weight being among the key points. The microstrip certainly saves quite a bit of weight, and this is a very neat application.

Broad-band applications, with a $1\frac{1}{2}$ or 2 : 1 band, are not readily catered for by waveguides, and if weight and compactness are featured, strip-line should be considered, provided that the performance required is not too high. Strip-line will normally give better than 2 : 1 standing-wave ratio over a broad band, and provided that one does not want stray coupling much better than 30 dB down, strip-line should be considered.

I have no doubt that there are a number of classified projects where extremely light weight, reproducibility and cheapness are important, and strip-line needs to be considered here.

Dr. Elizabeth Laverick: I was particularly interested in the adjustable short-circuit based on the idea of converting the microstrip line into waveguide beyond cut-off. I have used this type of plunger with success at 3 cm, and have also endeavoured to adapt it for use in low-Q-factor triplate line. The mechanical difficulties involved are even greater, for it is necessary to convert both halves of the triplate into waveguide beyond cut-off, and this entails sliding a metal strip over the centre conductor between the dielectric interfaces. On the other hand, in triplate line the energy is completely contained between the two ground planes, unlike microstrip, where energy is lost due to radiation. Hence in this respect the triplate short-circuit should be more efficient.

Were measurements made of the reflection coefficients of individual slots of the slot-coupled directional coupler? Over what dimensional limits does the single slot behave as a pure susceptance? Has the author any results on the match of the complete coupler over the waveband?

Mr. W. J. Bray: The salient advantages of strip-lines are that they are compact, light and offer the possibility of easy reproducibility at low cost. Perhaps their most serious disadvantage is their rather limited power-handling capacity, but others, less serious, arise from fringing-field effects and the difficulties of precise design of components.

Compactness and lightness suggest a possible field of application in airborne microwave equipment, notably on the receiver side; for pulse-radar equipment the power-handling limitation is a serious drawback, but there is a possible application in c.w. frequency-modulation radar equipment, where the power to be handled is relatively small.

On the ground, main-route microwave equipment offers limited possibilities. The high capacity equipment which provides several hundred telephone channels, or perhaps a television channel, uses equipment with conventional rectangular waveguide techniques, but the requirements for space are largely determined by the limited power efficiency of microwave equipment, e.g. 1–2 kW of mains power is used to radiate 1–2 watts from the aerial. Thus the space requirements are conditioned more by the power equipment and by its ancillaries than by the use of conventional rectangular-waveguide techniques.

However, detailed investigation might well be given to microwave equipment providing a small number of telephone channels direct to a subscriber's premises or for junction circuits. With microstrip techniques for the truly microwave part, printed-circuit techniques for the i.f. equipment, and compact forms of time-division multiplex equipment on printed circuits for the telephone channelling equipment, one can foresee a single unit, perhaps in the form of a cube of 18 in side with an 18 in diameter aerial attached to it, which would provide a small number of telephone channels at a relatively low cost in a form very convenient for use in the field over distances of up to 10 miles or so. However, it must be borne in mind that frequency space below 10 Gc/s would be difficult to find for this kind of application, and it would therefore be desirable to use frequencies above this. Would strip-line techniques, in fact, be suitable for applications above 10 Gc/s, such as that envisaged?

Mr. J. Hooper: If strip-line is to be used in the telephony application described by Mr. Bray, someone will be faced with

the tremendous problem of designing a repeater for a high-power system operating beside a low-power one.

Why does the author consider the parallel-line coupler to have such disadvantages? While the conductors must be very close together to ensure sufficient coupling, in view of the extreme accuracy claimed for the photo-etching technique, this would not seem to be a disadvantage.

So far, we have been given only insertion-loss measurements on the waveguide filters, and I should like to know something about the reflection coefficient over the pass band.

Mr. K. Foster: On the subject of measurement difficulties, I think that Mr. Dukes and Mr. Lewin speak from the aspect of microstrip mainly, for measurement is not so difficult with high-Q-factor strip-line. We have built a standing-wave indicator for this line, and its error appears to be about 2%. The main problem is that of setting the strip opposite the probe parallel with the probe travel; it has not yet been solved, and one must make a correction for this factor.

In the centre sections of the filter the branching arms appear to be very close together (I take it that this is high-Q-factor triplate). In our experience, this results in proximity coupling between the two arms. Has the author had any trouble with this in the design?

We employ an etching technique which is extremely simple. The part of the board to be used is painted with engineer's marking dye and the circuit is marked out. This region is then covered with transparent cellulose tape (preferably in one piece), which is cut along the marking lines and the surplus is peeled off. The surplus dye is then washed off with acetone, and the required copper circuit is left with a coating of dye and tape, the copper to be etched being uncovered. The board is then etched and the protective coatings are stripped and washed off.

Brigadier E. J. H. Moppett: The comment has been made that this technique is not yet greatly used in this country. Is this due to patent coverage?

Mr. F. J. H. Charman: The author points out the tendency among designers to think in terms of Zobel filters, and the fallacy of so doing. I agree, and wonder whether there may be the same tendency in measurement. Everyone is so accustomed to the idea of measuring standing waves with a slotted line that he does not necessarily automatically think of measuring them in any other way. Perhaps the difficulty lies there.

I have come to think in terms of measuring reflection coefficient instead of standing-wave ratios, and find that if one starts with reflection coefficient instead of impedances when teaching juniors, they understand much better. One can explain the Smith chart quite easily in terms of reflection coefficient, and having shown how that works, one can superimpose the impedance graticule and say, 'That is the transformation $Z = (1 + \rho)/(1 - \rho)$'.

I am leading on to the subject of measuring the reflection coefficient directly instead of deducing it from the standing-wave ratio. The author mentions its advantage in the matter of directional couplers and points to some of his latest work. This may lead to some fairly precise instruments; given these, I think that the measurement would be considerably simplified.

There is, of course, the alternative of the hybrid-ring type of directional device, but this is a resonant circuit. The directional coupler can be made a very wide-band affair, and in the v.h.f. range, at least, has been made to give a discrimination of over 60 dB at one frequency, or 40 dB over a 2 : 1 frequency band.

Mr. P. P. Eckersley: Have temperature tests been made and, if so, what is the result?

THE AUTHOR'S REPLY TO THE ABOVE DISCUSSION

Mr. J. M. C. Dukes (*in reply*): Mr. Lewin has endeavoured to put strip-transmission-line techniques in perspective with current waveguide practice. He has stressed, quite rightly, that it is not merely a matter of printing; additional piece-parts are required, which must be assembled to the printed board, and the completed circuit must be subjected to tests, the cost of which may form a substantial fraction of the whole. Nevertheless, a worthwhile economic advantage can occur in favourable applications. In the United States, for example, several firms are marketing individual components in enclosed cases at very competitive prices. Several such components have been tested, with rather good results. I agree with Mr. Lewin's comments concerning the attenuator, but as is made clear in the text, this instrument was intended originally for laboratory use only. Stray coupling can, of course, be reduced to any degree desired, but I agree that figures better than 30 dB are often difficult to achieve without structural complications, e.g. shielding. In regard to performance generally, it must be remembered, of course, that the effort so far devoted to these techniques is but a fraction of that spent on conventional waveguide circuits during the last two decades.

In answer to Dr. Laverick, I regret we did not make any extensive measurements of the reflection coefficients of individual slots, nor did we verify over what dimensional limits the single slot behaved as a pure susceptance. The reflection introduced by small slots was virtually negligible, as might be expected. The match of the complete coupler averaged 0·75, with a dip to 0·6 at one frequency, but this includes the input coaxial connector.

Mr. Bray asks whether strip lines are suitable for applications above 10 Gc/s; there is no fundamental objection, but practical application of the technique is bound to be more difficult, if only for mechanical reasons.

In reply to Mr. Hooper, the principal disadvantages of the simple parallel-line directional coupler are the complex character of the coupling mechanism and the tendency to resonate in the twin strip mode. Mr. Hooper and Mr. Foster have raised several questions concerning the re-entrant line filter. Except near the cut-off frequency, the input match was generally better than 0·7. Proximity coupling presumably does occur in the centre sections, but with little effect on the cut-off frequency. I have constructed a standing-wave indicator similar to that described by Mr. Foster and have encountered exactly the same difficulties. In our case the spacing between the ground planes was only $\frac{1}{8}$ in, which aggravated both this problem and also that of designing a probe having adequate rigidity and being free from spurious pick-up.

Brigadier Moppett is probably correct in assuming that the technique is not yet greatly used in this country. The principal deterrents have been the limited power-handling capacity, the absence of suitable materials and uncertainties in regard to performance.

In reply to Mr. Eckersley, the temperature limits quoted in Table 3 of Paper No. 2402 are adequate for most applications; a matter of rather more concern is the deterioration in power factor on exposure to extreme humidity.

621.376.56: 621.391

The Institution of Electrical Engineers
Paper No. 2454 R
Mar. 1958

A CODER FOR HALVING THE BANDWIDTH OF SIGNALS

By A. R. BILLINGS, B.Sc., Ph.D., Graduate.

(*The paper was first received 17th May, and in revised form 26th September,* 1957.)

SUMMARY

It is shown that it is possible to code a continuous message of finite bandwidth into a continuous signal of smaller bandwidth, provided that sufficient signal power is available. A possible practical system capable of performing this coding is described and it is shown that the exchange of bandwidth for power occurs in the same manner as for Shannon's ideal system.

(1) INTRODUCTION

It has become commonplace[1] to devise communication systems which allow the use of a reduced signal/noise ratio at the expense of increased bandwidth. In what follows, a system is described which reverses the process and reduces bandwidth at the expense of signal/noise ratio. This system is, in effect, a form of pulse code modulation (p.c.m.) in which the scale of numbers used to specify the code is increased, rather than decreased as in binary p.c.m.

(2) QUANTIZED PULSE SYSTEMS IN GENERAL

By popular usage, pulse code modulation has become associated with a system in which information is transmitted in binary form. This is rather misleading because any amplitude-quantized pulse system is, in fact, a pulse code modulation system. Thus, if a continuous waveform of bandwidth Δf is sampled the necessary $2\Delta f$ times per second, and the pulse amplitudes are quantized into n levels and then transmitted, the system is a pulse code modulation system using single-digit numbers in the scale of n to specify each sample amplitude.

In the binary p.c.m. system, the n levels are specified by binary numbers, and a positive or negative pulse is transmitted to specify each binary digit as either 1 or 0. Thus binary p.c.m. uses more pulses per second, and hence more bandwidth, than the p.c.m. system using a scale of n. The advantage gained in binary p.c.m. is that the transmitted power can be made smaller, since the receiver has only to distinguish between two levels instead of n levels, and in consequence the problem of regeneration is much simpler. There is, of course, no restriction upon the scale of numbers used to specify each quantized level, and tertiary and other p.c.m. systems have been devised which make use of other scales of numbers to obtain similar reductions of signal/noise level at the expense of bandwidth. This exchange is perfectly reasonable when designing wide-band modulation systems for use at ultra-high frequencies, where power is at a premium and bandwidth is virtually unlimited. However, when considering transmission through low-pass systems, such as telephone cables, some method of reducing bandwidth in exchange for increased transmitted power becomes desirable, provided that the exchange is an equitable one.

The technique employed to increase bandwidth is that of using a scale of numbers less than n. This suggests the technique to be used when reducing bandwidth, namely to use a scale of numbers in excess of n. The simplest method of doing this arises if the bandwidth is to be halved, in which case the scale of numbers must be increased to n^2. Two adjacent samples are then specified uniquely by a one-digit number in the scale of n^2, and only one code pulse is transmitted where two were previously required. There is, of course, no theoretical restriction on the amount of bandwidth division obtainable, and a reduction to a third could be obtained by specifying three adjacent samples by a one-digit number on the scale of n^3. Normally a reduction below about one-half will be prohibited by the large transmitted powers required. A simple illustration of coding from a scale of n to a scale of n^2 is provided by the representation of the number 23 in the scales of 5 and 25. If the letters A, B . . . Y represent numbers 0, 1 . . . 24, then in the scale of 5 the number is written ED whilst in the scale of 25 it is written X. This is because 23 can be represented either as $(4 \times 5^1 + 3 \times 5^0)$ or as (23×25^0).

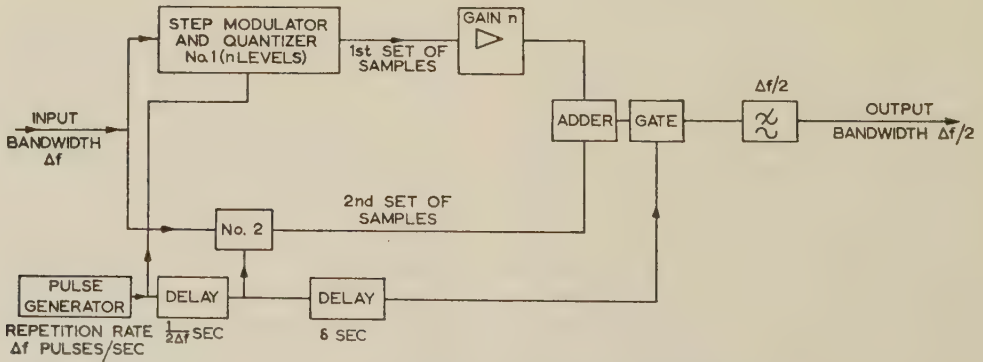


Fig. 1.—Form of coder which will code a signal of bandwidth Δf into a signal of bandwidth $\Delta f/2$.

(3) BANDWIDTH-HALVING SYSTEM

The main components of the bandwidth-halving system are what will be termed step modulators. A step modulator is a

Written contributions on papers published without being read at meetings are invited for consideration with a view to publication.

The paper is based on part of a thesis submitted for the degree of Doctor of Philosophy at London University.

Dr. Billings is in the Department of Electrical Engineering, University of Bristol.

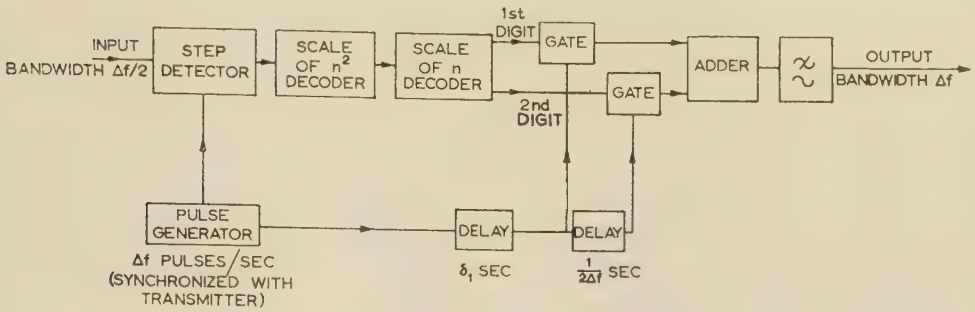


Fig. 2.—Device for decoding signal of bandwidth $\Delta f/2$ back into original signal.

evice which samples an input time function f_s times per second nd produces an output waveform which changes in amplitude t each sampling instant to that of the sampled waveform, and hich then has a constant amplitude until the next sampling nstant. Examples of step-modulated waveforms are shown in ig. 3.

The bandwith-halving system, shown in Fig. 1, is simple in rinciple, and will be described for the case where the distinguish- ble levels at the receiver are equal voltage intervals. The input ontinuous message of bandwidth Δf is supplied as input to two tep modulators, the outputs of which are amplitude-quantized nto n levels. The first step modulator is operated directly by ulses from a pulse generator having a repetition frequency Δf, nd the quantized amplitudes of the step wave produced are ermed the first step of samples, as shown in Fig. 3. The second tep modulator is operated at times $1/(2\Delta f)$ seconds after the rst, and the amplitudes are termed the second set of samples. he output of the first modulator is magnified by a factor n in stable amplifier and is added to the output of the second odulator. This sum is sampled at times δ seconds later than he second samples to provide what are called the composite amples, which are then transmitted as narrow pulses through a erfect low-pass filter of bandwidth $\Delta f/2$. The output of the lter is the continuous function which is sent through the trans- ission path to the receiver.

To regenerate the received continuous message the received ignal is sampled at a frequency Δf and each sample is labelled ith a number in the scale of n^2 and decoded into two numbers, ach in the scale of n, corresponding to the first and second amples. Pulses spaced $1/(2\Delta f)$ seconds apart with amplitudes roportional to these scale-of-n numbers are then formed and pplied to a perfect low-pass filter of bandwidth Δf. The output f this filter is a modified form of the original message, the modi- cation being that at the sample instants $1/(2\Delta f)$ seconds apart he received message is amplitude-quantized.

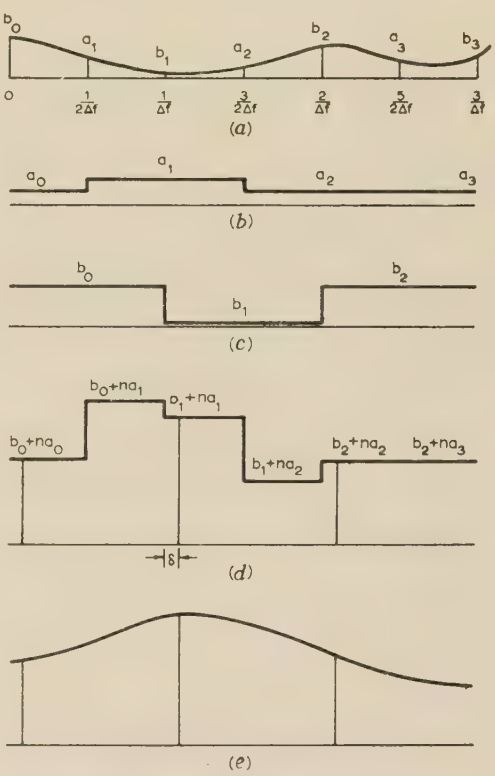


Fig. 3.—Waveforms produced in coding process.

(*a*) Input signal of bandwidth Δf and n levels sampled at frequency $2\Delta f$.
(*b*) Step-modulated form of first set of samples.
(*c*) Step-modulated form of second set of samples.
(*d*) Composite of amplified first samples and second samples (reduced amplitude scale).
(*e*) Transmitted signal of bandwidth $\Delta f/2$ and n^2 levels.

(4) THE RATE OF EXCHANGE OF BANDWIDTH FOR POWER

Shannon[2] has derived an expression for the capacity of a hannel for transmitting information when the bandwidth is Δf, he mean signal power is P_s and Gaussian noise of power P_n nterferes with the signal. This expression is

$$C = \Delta f \log_2 (1 + P_s/P_n) \text{ bits per second} \quad . \quad . \quad (1)$$

his expression is strictly only applicable to the ideal communica- on system, but it indicates how bandwidth and power should e exchanged in any good practical coding system. It will be hown that the coding system described above does exchange andwidth for power in this manner.

he maximum rate of transmission of information in the scale f n and scale of n^2 systems described above will be

$$R = 2\Delta f \log_2 n \text{ bits per second} \quad . \quad . \quad . \quad (2)$$

which is approached as the rate of incidence of decoding errors approaches zero. If the interfering noise powers in the scale of n and scale of n^2 systems are P_{n1} and P_{n2}, respectively, then, for the same incidence of decoding errors per digit, the ratio of the signal/noise ratios is[1]

$$\frac{P_{s2}/P_{n2}}{P_{s1}/P_{n1}} = \frac{(n^2-1)^2}{(n-1)^2} \quad . \quad . \quad . \quad . \quad . \quad . \quad (3)$$

which approximates to n^2 when n is large. This is in accordance with the ideal behaviour as predicted by Shannon's law. In the general case of a bandwidth reduction of $1/\alpha$,

$$\frac{P_{s\alpha}/P_{n\alpha}}{P_{s1}/P_{n1}} \simeq n^{2(\alpha-1)} \quad . \quad . \quad . \quad . \quad . \quad . \quad (4)$$

If the noise power is evenly distributed over the frequency band

and the power per unit bandwidth is the same for the compressed and uncompressed systems,

$$P_{s\alpha} = \frac{n^{2(\alpha-1)}}{\alpha} P_{s1} \quad . \quad . \quad . \quad . \quad . \quad (5)$$

(4.1) Possible Practical System

To illustrate the interchange of power and bandwidth, a possible practical system will be considered giving a bandwidth reduction of one-half. Let Δf be 3 kc/s and n be 100. Using formulae derived by Oliver, Pierce and Shannon[1] for the signal/noise ratios to give a probability of decoding error of 10^{-6}, the signal/noise ratios required for the scale of 100 and scale of 10 000 systems are 49·2 dB and 89·2 dB, respectively. If the noise is assumed to be 30 dB above thermal noise and the insertion loss between regenerative repeaters is 60 dB, the transmitted powers for the full- and half-bandwidth systems are 2 mW and 10 watts, respectively. This latter figure is a perfectly reasonable one, but, if the expense of twice the number of repeaters can be tolerated, the transmitted power in the half-bandwidth system can be reduced to 10 mW.

(5) COMMENT AND CRITICISMS OF BANDWIDTH-HALVING SYSTEM

It has been assumed that the distinguishable levels at the receiver are equal voltage intervals. If they are equal power levels, a square-root transfer device must be incorporated in the transmitter in front of the final low-pass filter. The purpose of this device is to convert the voltage scale into what McKay has called a proper scale.[3] Similarly a square-law device must be incorporated in the receiver.

One shortcoming of the system, which becomes important when working with signal/noise ratios near the threshold, is that the system is not an ambiguous index system.[4] This means that it is much more likely that drastic errors will be made in the interpretation of the second set of samples than in the interpretation of the first set. Another shortcoming is that perfect low-pass filters have been postulated. Such filters are not physically realizable, but can be approximated to if large delays can be tolerated. Such long delays are symptomatic of coding systems of this type.

It can therefore be concluded that continuous messages or signals which are amplitude-quantized can be coded into signals of smaller bandwidth without destroying information, and that the exchange of bandwidth for power is a reasonable one in the sense that it takes place as predicted by Shannon's law.

(6) BIBLIOGRAPHY

(1) Oliver, B. M., Pierce, J. E., and Shannon, C. E.: 'Philosophy of P.C.M.', *Proceedings of the Institute of Radio Engineers*, 1948, **36**, p. 1328.

(2) Shannon, C. E.: 'Communication in the Presence of Noise', *ibid.*, 1949, **37**, p. 10.

(3) McKay, D. M.: 'Quantal Aspects of Scientific Information', *Philosophical Magazine*, 1950, **41**, p. 284.

(4) Earp, C. W.: 'A Recent Development in Communication Technique', *Proceedings I.E.E.*, Paper No. 1328 R, July, 1952 (**99**, Part III, p. 181).

621.372.5

The Institution of Electrical Engineers
Paper No. 2464 R
Mar. 1958

AN APPROACH TO THE DESIGN OF CONSTANT-RESISTANCE AMPLITUDE EQUALIZER NETWORKS

By J. S. BELL, Graduate.

(*The paper was first received 8th July, and in revised form 9th October*, 1957.)

SUMMARY

It is often desirable to correct or to adjust the amplitude/frequency response of a network to meet certain stated requirements. Although it is usually not difficult to calculate the insertion loss of a correcting network, the inverse is rather more difficult. The paper suggests a method of designing constant-resistance amplitude equalizers to give a desired slope over the working range concerned. A typical example, involving the technique developed, is given in order to illustrate the method of adjusting the response of a velocity-type pick-up over part of the audio-frequency spectrum.

LIST OF PRINCIPAL SYMBOLS

N = Complex insertion-loss ratio.
ω_0 = Angular cut-off frequency, rad/s.
R_0 = Iterative resistance of network, ohms.
Z_1 = Impedance of shunt dual network, ohms.
Z_2 = Impedance of series dual network, ohms.
R_1 = Resistance associated with shunt dual network, ohms.
R_2 = Resistance associated with series dual network, ohms.
L_1 = Inductance associated with shunt dual network, henrys.
L_2 = Inductance associated with series dual network, henrys.
C_1 = Capacitance associated with shunt dual network, farads.
C_2 = Capacitance associated with series dual network, farads.

(1) INTRODUCTION

The subject of amplitude equalizers has already received considerable treatment,[1,2] and it is normally a comparatively simple matter to calculate the insertion loss of a given equalizer circuit. However, in practice, the problem usually occurs in the inverse form, and it is necessary to design a circuit to give a desired response. It is the purpose of this paper to suggest an easy and systematic approach to the problem.

Several graphs are given to facilitate the design of equalizers. Using the techniques outlined and referring to the example given, no great difficulty should be experienced with a practical problem. As for filters, owing to tolerances and strays it may be necessary to carry out final adjustments in the circuit itself should a particularly high degree of accuracy be desired. In most cases such final adjustments should not be found necessary.

(2) SPECIFICATION OF REQUIREMENTS

The design of filters and equalizing networks covers an extremely wide field, and their applications are both numerous and varied. It is therefore essential to decide the network required.

Equalization may, under certain conditions, be performed satisfactorily by the inclusion of either a capacitor or an inductor in an appropriate position. Frequently, however, the inclusion of such a simple device, or for that matter a simple T or π network is not satisfactory, especially where impedance matching is important over a wide frequency band. Any variation in impedance is particularly undesirable when measurements are to be taken. A further disadvantage is that, as both source and load impedance have to be taken into consideration and may vary from circuit to circuit, a comprehensive set of design curves is not easy to produce.

The chosen equalizer is therefore of the constant-resistance type and is easily regarded as a 4-pole network. Only one type is considered, i.e. a bridged-T network utilizing inductance, capacitance and resistance. This particular type of 4-pole network, if correctly matched, will present a sensibly constant resistance over a wide frequency range. This simplifies the design problem greatly, since, normally, equal source and load impedances can be arranged.

Following usual practice the amplitude response only will be considered; any phase change[3] may be considered or calculated if desired.

(3) FUNDAMENTAL APPROACH

The mathematical properties of the equalizer under consideration are closely allied to the theory of dual networks. The two dual networks considered are illustrated in Figs. 1(*a*) and 1(*b*).

Fig. 1.—The dual networks used in the formation of the bridged-T network.
(*a*) High-pass case. (*b*) Low-pass case.

Let us therefore construct a bridged-T network, using Fig. 1(*a*).

For constant resistance, Z_1Z_2 must be a constant, and this is chosen to be R_0^2, where R_0 is equal to the desired iterative impedance of the network.

Now
$$Z_2 = R_2 + j\omega L_2$$
$$\frac{1}{Z_1} = \frac{1}{R_1} + j\omega C_1$$

Satisfying the chosen condition,

$$\frac{R_2 + j\omega L_2}{\frac{1}{R_1} + j\omega C_1} = R_0^2$$

i.e.
$$R_2 + j\omega L_2 = \frac{R_0^2}{R_1} + j\omega C_1 R_0^2$$

Thus, separating real and imaginary parts,

$$R_1R_2 = R_0^2 \quad \ldots\ldots\ldots \quad (1)$$

$$\frac{L_2}{C_1} = R_0^2 \quad \ldots\ldots\ldots \quad (2)$$

Written contributions on papers published without being read at meetings are invited for consideration with a view to publication.
Mr. Bell is with Blackburn and General Aircraft, Ltd.

The complex insertion loss ratio is given by

$$N = 1 + \frac{R_0}{Z_2}$$
$$= \frac{R_0 + R_2 + j\omega L_2}{R_2 + j\omega L_2}$$
$$|N|^2 = \frac{(R_0 + R_2)^2 + \omega^2 L_2^2}{R_2^2 + \omega^2 L_2^2}$$

i.e. $$\text{Insertion loss} = 10 \log_{10}\left[\frac{(R_0 + R_2)^2 + \omega^2 L_2^2}{R_2^2 + \omega^2 L_2^2}\right] \quad . \quad (3)$$

It is easily seen that least attenuation occurs when $\omega \to \infty$.

Maximum attenuation occurs when $\omega = 0$.

$$\text{Maximum attenuation} = 10 \log_{10}\left[\frac{R_0 + R_2}{R_2}\right]^2$$
$$= 10 \log_{10} K^2, \text{ say}$$

Consider eqn. (3). Dividing through by R_2^2 and substituting,

$$\text{Insertion loss} = 10 \log_{10}\left[\frac{K^2 + \frac{\omega^2 L_2^2}{R_2^2}}{1 + \frac{\omega^2 L_2^2}{R_2^2}}\right] \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (4)$$

Let us define ω_0 such that

$$\omega_0^2 = K \frac{R_2^2}{L_2^2} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (5)$$

$$\text{Insertion loss} = 10 \log_{10}\left[\frac{K^2 + K}{1 + K}\right], \text{ where } \omega \to \omega_0$$
$$= 10 \log_{10} K \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (6)$$

Thus half of the maximum attenuation occurs when $\omega = \omega_0$. The corresponding frequency, f_0, will henceforward be called the cut-off frequency.

Similarly, by consideration of the dual network illustrated in Fig. 1(*b*), it may be shown that

$$\text{Insertion loss} = 10 \log_{10}\left[\frac{1 + \omega^2 C_2^2 (R_0 + R_2)^2}{1 + \omega^2 R_2^2 C_2^2}\right] \quad . \quad . \quad . \quad (7)$$

Eqns. (3) and (7) give the insertion loss for the high- and low-pass cases, respectively. The fundamental approach having now been covered, these two expressions will be examined in more detail.

(4) DERIVATION OF GRAPHS

Two expressions have now been obtained from which, given the relative component values, the insertion loss for either the high- or low-pass cases may be calculated at any desired frequency. Let us examine eqn. (3) in more detail.

It is seen that, if ω is doubled and at the same time L_2 is halved, the expression remains unaltered. Hence, if ω is moved in octaves, by suitable adjustment of L_2 a constant insertion loss will result. This is of significant practical value since the construction of attenuation Tables for the principal values of maximum attenuation is facilitated. Thus, by arranging to increase both the applied frequency, ω, and the frequency at which the attenuation is one-half its maximum value, ω_0, in octaves, it is a comparatively easy matter to construct a comprehensive Table for a particular maximum insertion loss. There is no need to reproduce one of these Tables, because the results are implied in any one of the curves shown in Fig. 3. Several

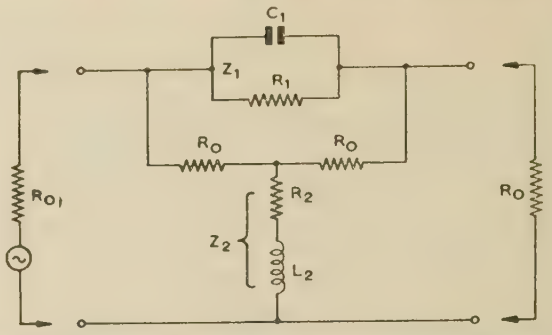


Fig. 2.—High-pass constant-resistance amplitude equalizer.

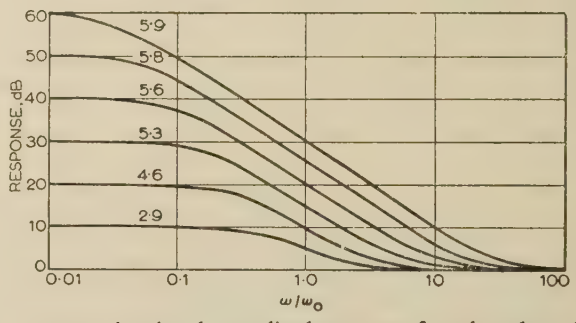


Fig. 3.—Curves showing the amplitude response for selected maximum attenuations.

Only the high-pass case is shown.
Note the differing slopes as indicated on the diagram.
The curves are plotted to a normalized frequency base.

curves for selected maximum attenuations plotted against a normalized cut-off-frequency base are illustrated.

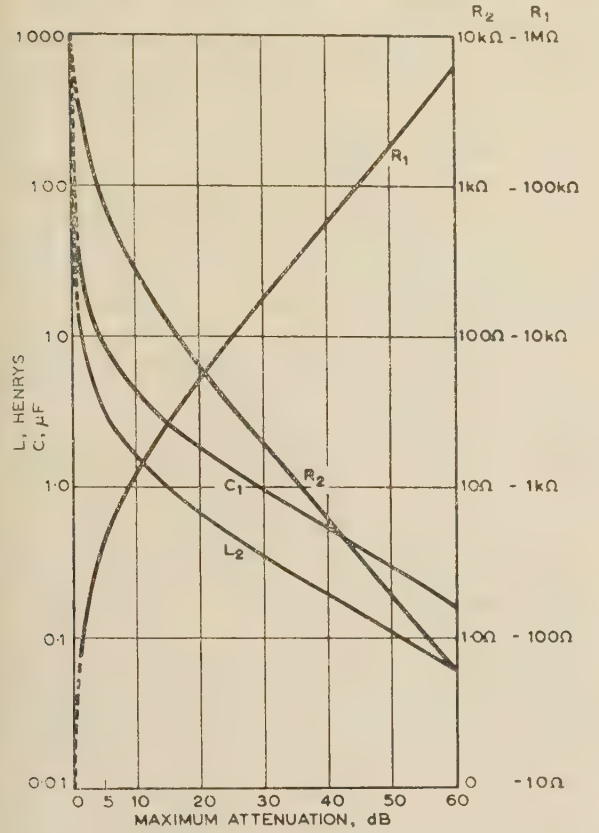
It will be appreciated that, in order to determine R_1, R_2, L_2, C_1 and K, use must be made of eqns. (1), (2), (5) and (6). Once these have been determined for a particular cut-off frequency, the appropriate values can be inserted in eqn. (3). (As the sums, products, divisions and logarithms of large numbers are involved, a desk calculating machine is used.) For normal degrees of accuracy such calculations will not be necessary; the graphs included with the text will prove sufficient.

Both the high- and low-pass cases for a particular maximum attenuation are identical but reversed. This is apparent since it will be remembered that, in the high-pass case, the attenuation is zero at infinite frequency, whilst in the low-pass case zero attenuation occurs at direct current. Thus, in forming a particular Table, it is not necessary to consider eqn. (7).

(5) INTERPRETATION OF RESULTS

In a similar manner, the corresponding Tables for any other desired maximum attenuation may be constructed. These are not shown because the corresponding values of Z_1 and Z_2 for both high- and low-pass cases are given graphically in Figs. 4(*a*) and 4(*b*), respectively. In neither case is it necessary to give a range of graphs for differing values of cut-off, because, having chosen to normalize the results from the start, it is an easy matter to scale up or down as desired. Hence only the values for a cut-off frequency of 50 c/s are designated. It is seen, therefore, that for any desired maximum insertion loss, the required values of the lumped components may be obtained.

Let us consider the family of graphs shown in Fig. 3. These represent the attenuation curves plotted against a normalized frequency base, and illustrate the differing slopes in decibels per octave obtained by selecting particular maximum attenuations. Each slope is sensibly constant over one decade, and, in addition,



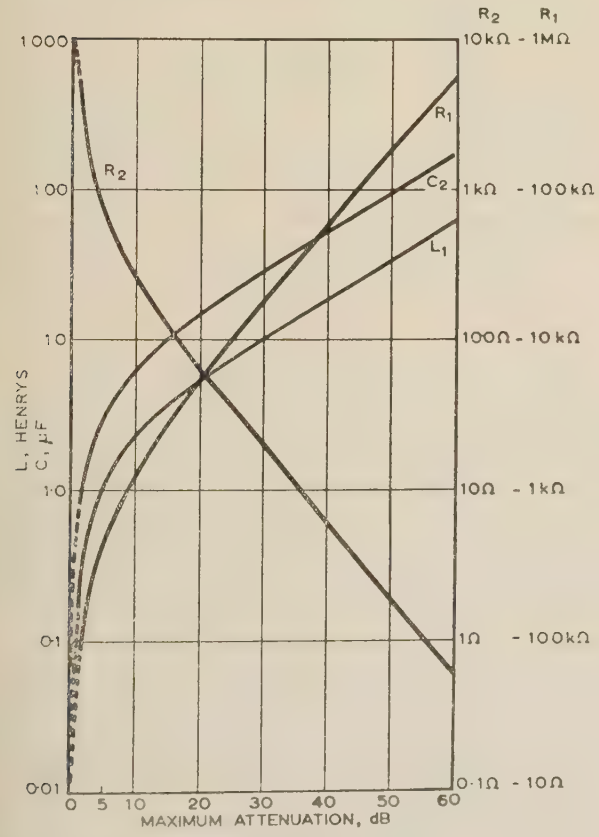


Fig. 4.—Graphs from which the component values are derived.
$f_0 = 50$ c/s. $R_0 = 600$ ohms.
(*a*) High-pass case. (*b*) Low-pass case.

Fig. 5.—Fundamental design graph illustrating the relationship between the slope desired and the degree of maximum attenuation required to produce it.

it is clear that, for any chosen curve, the slope will remain constant irrespective of the cut-off frequency.

By selecting further values of ultimate attenuation intermediate to those shown in Fig. 3, the graphical relationship between the slope, in decibels per octave, and the ultimate degree of attenuation may be plotted. This is the fundamental design graph and is illustrated in Fig. 5.

Thus it is possible to insert a desired loss (of either a positive- or negative-going slope) over a decade simply by selecting an equalizer with the relative maximum insertion loss and then arranging the appropriate cut-off frequency. To illustrate the principles involved, an example will be given.

(6) EXAMPLE

Let us suppose that an electro-mechanical device gives an output proportional to frequency as indicated in Fig. 6. It is desired that the response over the range 70–700 c/s should be flat, and also, if possible, that the resonant peak be removed or reduced in amplitude.

Upon measurement of the rate of fall-off, the loss will be found to be 1·2 dB/octave over the range concerned. Referring to Fig. 5, this will be found to correspond to a network having a maximum attenuation of 5 dB. By selecting the high-pass case with a cut-off frequency of 200 c/s and connecting it in tandem with the source (correct matching assumed), an 'equalized' response will result as shown. The values of R_1, R_2, C_1, and L_2 for the network concerned may be obtained from Fig. 4(*a*). These values are as follows:

$$Z_1 \begin{cases} R_1 = 470 \text{ ohms} \\ C_1 = 9{\cdot}1\,\mu\text{F} \end{cases} \qquad Z_2 \begin{cases} R_2 = 770 \text{ ohms} \\ L_2 = 3{\cdot}25 \text{ henrys} \end{cases}$$

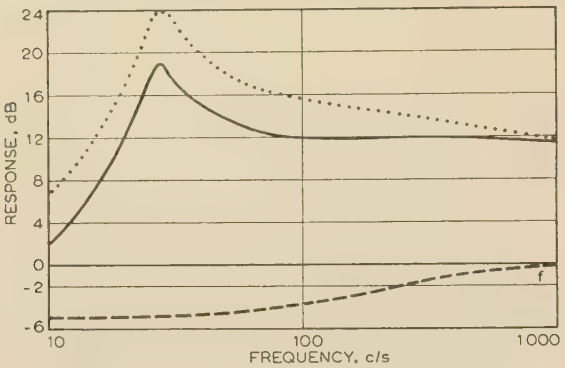


Fig. 6.—Correction of a velocity-type pick-up.

····· Unequalized output.
—— Equalized output.
---- Equalizer response.

(Had R_1, C_1, R_2 and L_2 been obtained by the rather lengthy procedure of calculation, after having referred to Fig. 5, their respective values would have been found to be 467 ohms, 9·09 μF, 771 ohms, and 3·27 henrys.)

Remembering that the above figures relate to a cut-off frequency of 50 c/s and not to the 200 c/s chosen, it is necessary to adjust C_1 and L_2 accordingly. As the required cut-off is two octaves higher their value must be halved twice. Thus the final values chosen for the correcting network are as follows:

$$Z_1\begin{cases} R_1 = 470 \text{ ohms} \\ C_1 = 2{\cdot}27\,\mu\text{F} \end{cases} \qquad Z_2\begin{cases} R_2 = 770 \text{ ohms} \\ L_2 = 0{\cdot}82 \text{ henry} \end{cases}$$

By selection of the same network, but with a cut-off frequency of 400 c/s ($C_1 = 1{\cdot}13\,\mu$F; $L_2 = 0{\cdot}41$ henry), equalization over the range 100–1 000 c/s will result.

(7) TRANSFORMATION OF ITERATIVE IMPEDANCE

All data and graphs relate to an iterative impedance of 600 ohms. Normally 600 ohms will prove suitable, especially as much test equipment is standardized at this impedance. Notes on impedance transformation are given in Section 11.

Should impedance transformation methods not be suitable in the particular case concerned, the design of an equalizing network having an iterative impedance different from 600 ohms presents no great difficulty using the data provided. The following relationships hold, where R_0' is the desired iterative impedance:

$$R_1' = R_1 \times \frac{R_0'}{600} \qquad L_2' = L_2 \times \frac{R_0'}{600}$$

$$R_2' = R_2 \times \frac{R_0'}{600} \qquad C_1' = C_1 \times \frac{600}{R_0'}$$

Hence, in the case of a required iterative impedance of 2 400 ohms, the design procedure is to assume a value of 600 ohms and then adjust the chosen values of the lumped components as indicated above.

Such a technique not only avoids the use of transformers or other devices but may well avoid serious losses. For instance, in using a transformer to match a source impedance of 60 kilohms to 600 ohms, a step-down ratio of 10 : 1 is unavoidable.

(8) CONCLUSION

Several of these networks have been designed utilizing the data given, and extremely good results have been achieved. The primary object of correcting the response obtained from a transducer, over a desired frequency range, has been fulfilled.

No filter or equalizing network is capable of undertaking all tasks set before it, and this type of equalizer has, besides its assets, its limitations. Provided that it is sensibly used, however, excellent results can be obtained.

As the network considered is a simple *RC* or *RL* circuit the maximum slope of the attenuation curve is 6 dB/octave. Should higher slopes be required, tandem connections are called for, and as the impedances of correctly designed networks are constant, such connections will not present any difficulty. The effective frequency range covers only one decade.

At lower maximum attenuations [Fig. 4(*a*)] and higher maximum attenuations [Fig. 4(*b*)] the capacitances and inductances become unwieldy, though they improve as the cut-off frequency increases. Two methods are available to avoid this difficulty. Impedance transformation may be resorted to, but as transformers and/or valves may be required, a better method is to employ two or more networks in tandem, thus giving the desired results using sensible values of components.

Essentially, the circuits described are designed to work over the audio-frequency range. At higher frequencies the stray capacitances become important, and careful design is required to avoid unwanted effects.

In the case considered it must be remembered that equalization is only obtained by inserting appropriate losses. Hence, in certain applications provision may have to be made, e.g. by employing an amplifying stage, to make good the losses incurred.

(9) ACKNOWLEDGMENT

The author wishes to thank Dr. H. Fuchs, Chief Electronics Engineer of the Blackburn and General Aircraft, Ltd., for permission to publish the paper.

(10) REFERENCES

(1) TERMAN, F. E.: 'Radio Engineers' Handbook' (McGraw-Hill, 1943).
(2) SANDEMAN, E. K.: 'Radio Engineering, Vol. 2' (Chapman and Hall, 1947).
(3) GRIFFITHS, J. W. R., and MOLE, J. H.: *Electronic and Radio Engineer*, 1957, **34**, p. 26.
(4) 'Output Transformerless Amplifiers—A General Review', *Wireless World*, 1957, **63**, p. 58.

(11) APPENDIX

Although the paper is applicable to the equalization of lines and similar networks, it was originally prepared with the object of correcting or adjusting the response of crystal accelerometers and velocity-type pick-ups. It is essential with crystal accelerometers, and sometimes desirable with velocity-type pick-ups, to load the transducer with a cathode-follower stage. The following Sections give a brief summary of the essential details of the cathode-follower method and also the more popular methods of impedance transformation which may prove useful in such work.

(11.1) Cathode-Follower Methods

Neglecting stray capacitance, the input impedance of a cathode-follower stage may easily be derived by standard methods. Difficulties sometimes arise when loading the cathode-follower and then calculating its output impedance. An approach is suggested to those who are not familiar with the problem.

By suitable analysis, the internal resistance of a cathode-follower may be shown to be $r_a/(\mu + 1)$, and under normal circumstances this approximates to $1/g_m$. The further approximation that the output impedance of a cathode-follower is

$1/g_m$ is only true when the load is high compared with $1/g_m$, since it is their parallel combination which must be considered. It follows directly that, in ascertaining the output impedance of a cathode-follower stage, the effect of the load must be considered. Let us therefore consider some of the more usual methods. Correct bias, or an adequately decoupled bias resistor, is assumed in all cases.

(11.1.1) **Load-Resistor Case.**

If the cathode-follower is loaded with a resistance R_k, which includes all effective resistance between cathode and earth, the resulting output impedance is the parallel combination of $1/g_m$ and R_k. Thus the values chosen for both $1/g_m$ and R_k are deciding factors in determining the output impedance. In calculating the gain of such a stage, care must be taken in the determination of the effective cathode load R_k', as any load connected will appear across R_k and will modify the effective load presented to the valve. Hence, in practice, the total load presented should be determined first.

(11.1.2) **Loaded with Equalizer or Line.**

If the cathode-follower is loaded with a constant-resistance network or transmission line (assuming a d.c. path) and the networks are correctly terminated, the load presented to the cathode-follower will be the iterative or characteristic impedance, respectively. Conversely, the load presented by the valve to the equalizer or line will be approximately $1/g_m$ as before. The design procedure is to arrange that $1/g_m = R_0$ or Z_0, respectively. The stage gain for perfect matching is just under one-half.

(11.1.3) **Transformer Coupling.**

Should a d.c. path not be available through the equalizer or line, and provided that capacitive coupling introduces too much loss, a further method is to use transformer coupling. However, unless care is taken in design the voltage gain may be very low indeed. The valve will be presented with a load as given by eqn. (8), and if an attempt is made to increase this load by means of a step-down transformer, two things occur simultaneously. First a step-down loss is introduced and secondly padding in the transformer secondary circuit (introducing further voltage loss) is necessary in order to match the loading network to the now transformed impedance of the valve. In order to avoid such difficulties a turns ratio of about unity should be chosen.

(11.1.4) **Push-Pull Output.**

To avoid transformer coupling and its attendant difficulties, two cathode-follower stages may be operated under push-pull conditions.[4]

(11.2) Transformer Method

Perhaps the easiest way of matching is by the use of a simple transformer where the turns ratio has been chosen to give the desired change in impedance. A disadvantage of the method is that, when increasing the apparent load, the voltage transferred is divided by the turns ratio chosen. Care must be exercised in the design of the transformer so as to ensure that it has an even response over the frequency range concerned. Special precautions are necessary at low frequencies in order to avoid any undesirable effects.

From elementary theory, the following general relationship holds:

$$Z_{in} = Z_p + k^2(Z_s + Z_L) \quad . \quad . \quad . \quad . \quad . \quad (8)$$

where Z_{in} = Input impedance presented by the transformer primary winding.
k = Step-down ratio = (Number of primary turns)/(Number of secondary turns).
Z_s = Impedance of secondary winding.
Z_L = Load impedance.
Z_p = Primary impedance.

(11.3) Other Methods

The two most obvious cases have been considered, but there is no reason why any other method should not be used, provided that matching is achieved over the frequency range concerned. For instance, should gain be required, a conventional amplifier stage may be of particular use for networks having a high iterative impedance.

621.385.2

The Institution of Electrical Engineers
Paper No. 2493 R
Mar. 1958

SPACE CHARGE AS A SOURCE OF FLICKER EFFECT

By C. S. BULL, Ph.D., Associate Member.

(*The paper was first received 24th November, 1956, in revised form 20th August, and in final form 5th November, 1957.*)

SUMMARY

The capacitive effect of space charge in a space-charge-limited or saturated diode is well known, but it has not been taken into account in forming a theory of the fluctuations.

The capacitive current evoked from the space charge by the fluctuations in anode voltage consists of the motion of free electrons and is therefore itself subject to fluctuations. By including these it is found that three distinct types of fluctuation exist together. One is the shot noise, depending in the well-known way on the effect of space charge in controlling the anode current. The second is an enhancement of the shot noise by an amount dependent on the ratio between the capacitance due to space charge and the total circuit capacitance. It is like the shot noise in that the total mean-square voltage is proportional to the resistance of the anode circuit. It is, however, unlike the shot noise in that it varies with anode voltage even when the valve is saturated. The third is a flicker effect dependent on the magnitude of the electronic capacitance, and not requiring the postulation of a life-time of metastable states on the electrodes.

In the course of the paper a general form of Campbell's theorem is given, together with a discussion of the way in which the theoretical considerations must be made to match the experimental conditions. Also, new probability generating functions are put forward and utilized.

(1) INTRODUCTION

The capacitive effect of space charge has been known for a long time. It is only recently, however, that the proper way of calculating it has been found[1] and applied to the simplest possible valve, a diode having zero emission velocity. The value of the capacitance turns out to be

$$C_e = \frac{\partial Q_s}{\partial V} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (1)$$

where Q_s is the total space charge and V is the anode voltage. In general there are other components of the capacitance due to charges bound on the electrodes, but we shall be concerned only with that part, C_e, which is due to free electrons in the inter-electrode space.

Experimental results obtained by Mr. Cotterhill, not yet published but kindly shown to the author, show that a capacitive effect due to space charge of the same magnitude as C_e arises in valves having appreciable emission velocities. C_e is comparable with the static (cold) capacitance of the electrodes in magnitude but is negative or positive, according to whether the valve is space-charge limited or not. Eqn. (1) can therefore be taken as a satisfactory starting-point for an elementary theory of the fluctuations due to space charge.

In our theory we must remember that the bandwidth of the amplifier used for measuring the fluctuations is of necessity limited, so that we cannot record the passage of single electrons. We can record only the effect of groups of r electrons passing in successive intervals of time, τ, given by

$$\tau = \frac{1}{f_b} = RC \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (2)$$

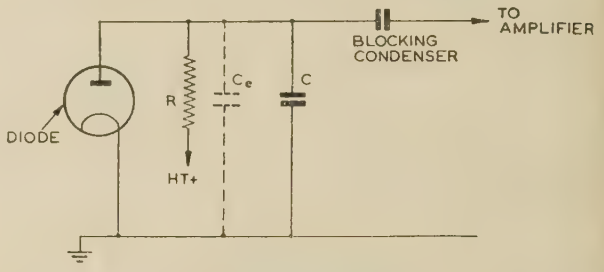


Fig. 1.—The diode circuit.

$C_e = \partial Q_s/\partial V$ represents the capacitive effect of the space charge. It is assumed to be much smaller than the total circuit capacitance, C.

where f_b is the bandwidth of the anode circuit and R and C are its resistance and capacitance. Fig. 1 illustrates the very simple circuit under consideration. Eqn. (2) is used as a convenient approximation to the relation between τ and f_b. A more exact expression would complicate the issue without yielding greater physical insight, and could at most change the calculated results by an unimportant amount without altering their nature.

It will be noticed also that we consider only the time-constant or the total bandwidth f_b of the anode circuit, and do not use the elements of bandwidth so frequently brought into consideration in studying the fluctuations. The element of bandwidth or unit bandwidth is certainly a mathematical fiction. In what follows we shall as far as possible avoid such fictions, especially in situations in which they can confuse or vitiate the calculations, and adhere to mathematical operations corresponding to the physical actions taking place and the measurements being made.

Also, it must be remembered that there is invariably a blocking condenser between the valve under test and the amplifier. Consequently the measurement consists in recording successive increments of a voltage δV_1 due to the change from n electrons passing during one interval of time τ to m during the next. δV_1 is given by

$$\delta V_1 = (n - m)\frac{e}{C} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (3)$$

The numbers n and m fluctuate independently. They must be integral; otherwise the number passing in one period would be connected with that passing during previous and successive intervals. The range of possible values of n and m is determined by two identical probability distributions, but their actual values are not known during any particular period τ.

A study of the fluctuations in δV_1 will give us the shot-noise voltage, the probability distributions given to n and m depending on whether or not the valve is space-charge limited.

The voltage δV_1 evokes a charge

$$\delta Q_s = \frac{\partial Q_s}{\partial V}\delta V_1 = C_e \delta V_1 \quad . \quad . \quad . \quad . \quad (4)$$

from the space charge, which, when it appears on the circuit capacitance C develops an additional voltage δV_2 given by

$$\delta V_2 = \frac{1}{C}\frac{\partial Q_s}{\partial V}\delta V_1 = \frac{C_e}{C}\delta V_1 \quad . \quad . \quad . \quad . \quad (4a)$$

Written contributions on papers published without being read at meetings are invited for consideration with a view to publication.

Dr. Bull is Reader in Physics (Electronics) at the College of Technology, Birmingham.

The total resulting anode voltage is then

$$\Delta V = \delta V_1 + \delta V_2 = (n - m)(1 + p)\frac{e}{C} \quad . \quad . \quad . \quad (5)$$

where $p = C_e/C$.

Now p is related to the differential coefficient $\partial Q_s/\partial V$, which is one of the characteristics of the valve. Mathematically a differential coefficient is a limiting value attained as the increments approach zero. However, it has been shown in a previous paper[2] that such differential coefficients have no constant value for voltages such as δV_1, which are as small as the fluctuations. From physical considerations the graph of Fig. 2 could be drawn

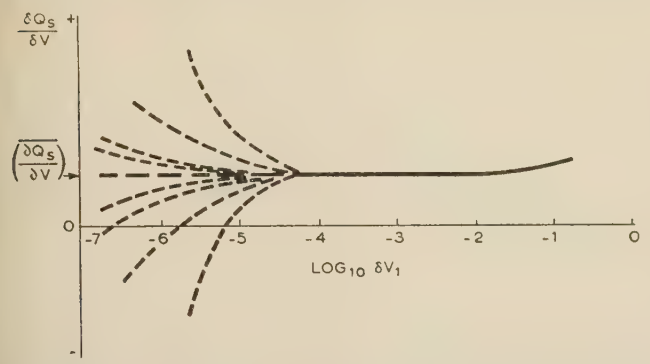


Fig. 2.—Showing how the quotient $\delta Q_s/\delta V$ approaches the differential coefficient $\partial Q_s/\partial V$ as δQ_s and δV are reduced, and becomes of unknown value when δV is comparable with or less than the fluctuation.

showing how, when very small voltages are concerned, the differential coefficient becomes blurred, fluctuating possibly from positive to negative values.

We are not permitted, however, to put any value we please for p. One restriction is that the total charge making up the voltage ΔV must be an integral number of electrons. This could be achieved by taking particular fractional values for p, such as $1/(n - m)$; $2/(n - m)$, . . . etc. Such a set of values of p would, however, connect the values of n and m, so that they could not fluctuate independently. We are obliged, therefore, to take for p only integral values.

The value of p is dependent on two factors:

(*a*) The state of the space charge during the period in which the p electrons are liberated.

(*b*) The necessity that the average value of p should be equal to the ratio C_e/C.

The simplest way to satisfy condition (*a*) is to make p proportional to n, i.e. to put

$$p = nk \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (6)$$

If now n and k are to fluctuate independently, n is integral and their product p is also integral, then k must itself be an independently fluctuating integer. We can now write an expression for the instantaneous voltage ΔV:

$$\Delta V = (n - m)(1 + nk)\frac{e}{C} = r\frac{e}{C} \text{ (say)} \quad . \quad . \quad (7)$$

in the study of which n, m, k are three independently fluctuating integral numbers. We shall now outline very briefly the methods we shall use.

(2) PROBABILITY GENERATING FUNCTIONS

A probability generating function (p.g.f.) for integral numbers is a power series of which the coefficients, a_n, of the nth power of the variable, i.e. x^n, is the probability that n occurs. The p.g.f. describing n is therefore

$$N(x) = \sum_n a_n \,.\, x^n \text{ with } N(1) = \sum_n a_n = 1$$

also $$N'(x) = \sum_n n a_n \,.\, x^n \text{ so that } N'(1) = \sum_n n a_n = \bar{n}$$

and $$N''(x) = \sum_n n(n-1) a_n \,.\, x^{n-2} \text{ so that}$$

$$N''(1) = \overline{n^2} - \bar{n}$$

The mean-square deviation of n is

$$\sigma_n^2 = \overline{(n - \bar{n})^2} = N''(1) - N'(1)^2 + N'(1)$$

No further properties of the p.g.f.'s will be required, but we shall have to find the p.g.f.'s of various sums and products of numbers. These will merely be stated without proof. The terminology used will indicate the combinations involved:

$(N + M)(x) = N(x) \,.\, M(x)$ is the p.g.f. for $(n + m)$

$(N - M)(x) = N(x) \,.\, M(x^{-1})$ is the p.g.f. for $(n - m)$

$(N \,.\, K)(x) = N(K(x))$ is the p.g.f. for (nk)

$(M \,.\, N \,.\, K)(x) = M(N(K(x)))$ is the p.g.f. for (mnk)

$(1 + P)(x) = xP(x)$ is the p.g.f. for $(1 + p)$

Using these forms we find that the p.g.f. for the integer r is

$$R(x) = (N - M)(xP(x))$$

$$= N(xP(x)) \,.\, M\left(\frac{1}{xP(x)}\right)$$

which demonstrates that a p.g.f. for r can be constructed. It is, however, a complicated function, and we shall make our problem easier by dividing it into three terms corresponding to the three simultaneous processes taking place:

$$r = (n - m) + n \,.\, nk - m \,.\, n \,.\, k$$

and we shall require the variances

$$\sigma_r^2 = \sigma_{n-m}^2 + \sigma_{n \,.\, nk}^2 + \sigma_{m \,.\, n \,.\, k}^2$$

The p.g.f. for $n \,.\, nk$ can be written down, but it is simpler to replace the first n by $\bar{n}$, since $\bar{n}$ is so large that its fluctuation will be relatively negligible; thus

$$\sigma_r^2 = \sigma_{n-m}^2 + \sigma_{\bar{n} \,.\, nk}^2 + \sigma_{m \,.\, n \,.\, k}^2 \quad . \quad . \quad . \quad . \quad (8)$$

(3) A GENERAL THEOREM FOR THE TOTAL MEAN-SQUARE VOLTAGE

Campbell's theorem can be expressed by the equation

$$\overline{V_t^2} = f_b \bar{r} \int_0^\infty V(t)^2 dt$$

where $V(t)$ is the voltage, expressed as a function of time, set up by the passage of one electron. The product $f_b \bar{r}$ is the average number of electrons passing per second and $\overline{V_t^2}$ is the total mean-square voltage set up by the random passage of the electrons.

The derivation of this theorem[3–6] has been based on the mathematical concept of dividing time into such small elements that the passage of one electron during one element is a rare event, while the probability that two electrons pass is so small that it may be neglected.

Now, for a current of 1 mA there are about 10^{16} electrons per second, so that the elements of time in this consideration are at least two orders smaller than 10^{-16} sec, which, according to the

Heisenberg uncertainty principle, will give an uncertainty in the energy of the electrons of about 10^3 eV. But we know the energy of the electrons in ordinary valves much more accurately than that. From this procedure we are very likely, therefore, to obtain results which have little physical significance. This is in fact seen to be the case if Campbell's theorem is applied to two valves, one having a saturation current and the other a space-charge-limited current of the same amount, while the circuit time-constants and times of flight of the electrons are arranged to be the same. It is well known that the fluctuations will be very different, while with Campbell's theorem we cannot distinguish between the valves.

We shall circumvent this difficulty by stating without proof a new and more general theorem:

$$\overline{V_t^2} = f_b \overline{r^2} \int_0^\infty V(t)^2 dt \quad . \quad . \quad . \quad . \quad . \quad (9)$$

This equation can be obtained very readily on the assumption that the amplifier is not able to record the passage of each individual electron, but only the combined effect of the passage of r electrons f_b times per second. The voltage recorded in every period τ is $rV(t)$, and from this the energy dissipation can be calculated. Eqn. (9) can then readily be deduced without introducing anti-physical assumptions.

Eqn. (9) can be rewritten

$$\overline{V_t^2} = f_b\left(\bar{r}^2 + \sigma_r^2\right) \int_0^\infty V(t)^2 dt \quad . \quad . \quad . \quad . \quad (9a)$$

and, by inserting for r a Poisson distribution and neglecting the effect of the average anode current, we obtain as a special case of eqns. (9) and (9a)

$$\overline{V_t^2} = f_b \bar{r} \int_0^\infty V(t)^2 dt \quad . \quad . \quad . \quad . \quad . \quad (9b)$$

which is Campbell's theorem. It has now, however, been placed on a sound physical basis, and its limitations are more clearly seen.

For our particular problem we are concerned only with cases for which $(\overline{n - m}) = 0$, so that $\bar{r}$ is always zero, and we shall use the new theorem in the form

$$\overline{V_t^2} = f_b \sigma_r^2 \int_0^\infty V(t)^2 dt \quad . \quad . \quad . \quad . \quad . \quad (9c)$$

We also note that since $V(t) = (e/C) \exp(-t/RC)$, the integral in all cases is $e^2R/2C$, and

$$\overline{V_t^2} = \frac{f_b \sigma_r^2 e^2 R}{2C} \quad . \quad . \quad . \quad . \quad . \quad . \quad (10)$$

We shall now consider the separate terms in eqn. (8) which make up the value of σ_r^2.

(4) THE SHOT NOISE

The shot noise, whether or not the current is space-charge limited, is given by the first term of eqn. (8), and for it we have

$$\overline{V_{n-m}^2} = \frac{f_b \sigma_{n-m}^2 e^2 R}{2C}$$

Using the p.g.f. for $(n - m)$ given in Section 2 and remembering that $N(x) = M(x)$, it is found that

$$\sigma_{n-m}^2 = \sigma_n^2 + \sigma_m^2 = 2\sigma_n^2$$

But the value of m for one period becomes the value of n for the next. We have therefore counted each value of n twice too often, so that for our present purposes we must put

$$\sigma_{n-m}^2 = \sigma_n^2 \quad . \quad . \quad . \quad . \quad . \quad . \quad (11)$$

from which we get

$$\overline{V_{n-m}^2} = \frac{f_b \sigma_n^2 e^2 R}{2C}$$

Experience has shown that, for a saturation current, the distributions of n and m are Poisson distributions, for which

$$N(x) \equiv M(x) = \exp \bar{n}(x - 1)$$

If we call Γ^2 the space-charge smoothing factor and σ_{nP}^2 the variance for a Poisson distribution of n, then

$$\sigma_n^2 = \Gamma^2 \sigma_{nP}^2 = \Gamma^2 \bar{n}$$

so that, finally, since the anode current I is equal to $f_b \bar{n} e$,

$$\overline{V_{n-m}^2} = \Gamma^2 \frac{eIR}{2C}$$

Putting $\Gamma^2 = 1$ we get Schottky's original expression for the noise of a saturation current. The space-charge smoothing factor now appears, not as an 'ordering' of the electrons, but as a change in the probability distribution when space charge is present.

Schottky's expression can also be obtained if it is assumed that a fluctuation current $2eI$ per unit bandwidth flows through the combined resistance and capacitance in parallel, and integrating the mean-square voltage over a very large band of frequencies. Consequently, mean-square voltages such as $\overline{V_{n-m}^2}$ which are proportional to R are like the shot noise in being independent of the bandwidth in the sense that they can be represented by a constant current per unit bandwidth. In later Sections we shall obtain mean-square fluctuation voltages which are not proportional to and increase more rapidly than R. Consequently they can be represented by a current per unit bandwidth which increases as the bandwidth is decreased, and may be classified as a manifestation of the flicker effect. Before we examine them we shall study the effect of the space charge in enhancing the shot noise.

(5) THE ENHANCED SHOT NOISE

To calculate the value of $\sigma_{\bar{n}.nk}^2$ we require the p.g.f. $(N \,.\, K)(x) = N(K(x))$. Using the processes indicated in Section 2 it is readily found that

$$\sigma_{nk}^2 = \bar{n}\sigma_k^2 + \sigma_n^2 \overline{k}^2 \quad . \quad . \quad . \quad . \quad . \quad (12)$$

This variance occurs for each of the $\bar{n}$ electrons we have assumed to pass in the period τ. Consequently for the $\bar{n}$ electrons passing simultaneously so far as the measuring amplifier is concerned, the total variance will be $\bar{n}$ times greater, so that

$$\sigma_{\bar{n}.nk}^2 = \bar{n}\sigma_{nk}^2 = \overline{n^2}\sigma_k^2 + \bar{n}\sigma_n^2 \overline{k}^2 \quad . \quad . \quad . \quad . \quad (13)$$

$$= \overline{n^2}\sigma_k^2 + \Gamma^2 \,.\, \overline{n^2}\,\overline{k}^2$$

$$= \overline{n^2}\sigma_k^2 + \Gamma^2\left(\frac{C_e}{C}\right)^2 \quad . \quad . \quad . \quad . \quad . \quad (13a)$$

Similarly, for the study of the product mnk we use the p.g.f.

$(M . N . K)(x) = M(N(K(x)))$ and it is straightforward to find that

$$\sigma^2_{mnk} = \sigma^2_{\bar{n}}\left(\frac{C_e}{C}\right)^2 + \sigma^2_{\bar{n}}\left(\frac{C_e}{C}\right)\bar{k} + \bar{n}\,.\,\bar{n}\sigma^2_k \quad . \quad . \quad (14)$$

The expressions for both $\sigma^2_{\bar{n}.nk}$ and σ^2_{mnk} contain terms in $\bar{n}$ and C_e/C, which, on being added together, give expressions which are of the shot noise type, proportional to R, and which we shall call an 'enhanced shot noise':

$$\overline{V^2_{eh}} = \left[2\Gamma^2\left(\frac{C_e}{C}\right)^2 + \Gamma^2\frac{C_e}{C}\bar{k}\right]\frac{\bar{n}e^2R}{2C}$$

$$= \Gamma^2\frac{eIR}{2C}\left[2\left(\frac{C_e}{C}\right)^2 + \frac{C_e}{C}\bar{k}\right] \quad . \quad . \quad . \quad (15)$$

Two distinct shot-noise effects have now been given, one due to the initial fluctuation of the anode current, and dependent on whether or not the current is space-charge limited, being smaller when the current is so limited. The other is an addition to the first, dependent on the extent to which the capacitance is influenced by the presence of space charge. Space charge exists, of course, even at the saturation current, especially when the anode voltage is only slightly greater than the saturation voltage, so that the enhancement of the fluctuations could lead to a shot noise greater than the full shot noise described by Schottky, over certain regions of the characteristic. The ratio C_e/C is not necessarily negligible, but it would be unwise in view of the approximate nature of the theory being developed to attempt a firm estimate of its value. However, it has been pointed out[2] that a considerable discrepancy exists between the theoretical estimates of the fluctuation of space-charge-limited currents and the observed fluctuations, and a factor independent of Γ^2 and anode current, such as C_e/C, would be of the right kind to explain the discrepancy.

(6) THE FLICKER EFFECT

Looking again at eqns. (13) and (14), we find in each a term $\bar{n}\,.\,\bar{n}\sigma^2_k$, from which we get a total mean-square voltage

$$\overline{V^2_f} = 2\frac{f_b\bar{n}\bar{n}\sigma^2_k e^2R}{2C} \quad . \quad . \quad . \quad . \quad . \quad (16)$$

$$= 2\frac{eIR}{2C}\sigma^2_k\bar{n}$$

We can make progress in the interpretation of this component of the fluctuation by investigating the distribution of the values of k. We have seen that $\partial Q_s/\partial V$ fluctuates from positive to negative; n, however, is always positive, so that k must fluctuate from positive to negative. This condition, together with the condition that k must be integral, leads us to suggest a p.g.f. for k having only three terms, representing the possibilities that k may be $+1$, 0 or -1, namely the trinomial

$$K(x) = ax + b + cx^{-1} \quad . \quad . \quad . \quad . \quad (17)$$

From this it can readily be seen that

$$\bar{k} = a - c$$

and

$$\sigma^2_k = (a + c) - (a - c)^2$$

This last expression is essentially positive, as it should be, since a, b and c are all fractional and $(a + b + c) = 1$.

As $\bar{n}$ is increased by increasing the resistance R, thereby decreasing the bandwidth, $\bar{k}$ decreases in such a way as to keep

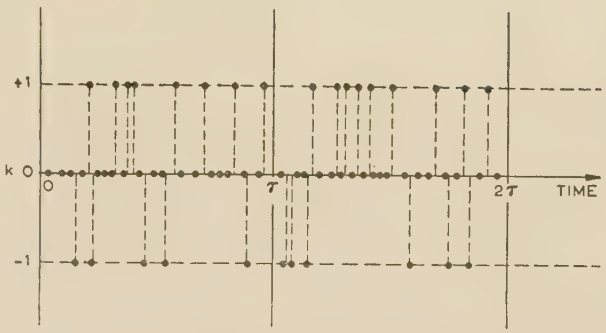


Fig. 3.—Fluctuation of the integer k.

Each dot on the lines $k = +1, 0, -1$ represents the passage of one of the n electrons during each interval τ.

the product $\overline{nk}$ $(= C_e/C)$ constant. It does not follow, however, that a and c should decrease as k decreases; it is necessary only that they should become more and more nearly equal. Fig. 3 shows a possible set of values of k over two successive periods τ. It will be apparent that it is physically unlikely that, as $\bar{k}$ decreases to zero, a and c will change appreciably, since the events in the inter-electrode space will not be critically dependent on $\bar{k}$. We can therefore take σ^2_k to be very nearly constant independently of $\bar{k}$, $\bar{n}$ or τ. Remembering now that $f_b\bar{n}e = I$, we can reduce the expression for $\overline{V^2_f}$ to

$$\overline{V^2_f} = I^2R^2\sigma^2_k$$

This fluctuation voltage increases more rapidly with R than do the expressions for the two kinds of shot noise, so we can infer that as the bandwidth is decreased the fluctuation expressed in the usual way as a current per unit bandwidth increases. This behaviour is in contrast with that of the shot noise, and is a feature of what has hitherto been called the flicker effect. We have here, however, obtained this result without assuming the existence of metastable states on the electrodes, as has always previously been done.

(7) DISCUSSION

We can now collect the expressions for the three types of noise contributed simultaneously by the fluctuations, getting a total mean-square voltage

$$\overline{V^2_t} = \overline{V^2_{n-m}} + \overline{V^2_{\bar{n}\,.\,nk}} + \overline{V^2_{mnk}}$$

$$= \overline{V^2_{n-m}} + \overline{V^2_{eh}} + \overline{V^2_f}$$

$$= \frac{eIR}{2C}\left\{\Gamma^2\left[1 + 2\left(\frac{C_e}{C}\right)^2 + \frac{C_e}{C}\bar{k}\right] + 2\bar{n}\sigma^2_k\right\}$$

In the second expression for $\overline{V^2_t}$, the terms $\overline{V^2_{nnk}}$ and $\overline{V^2_{mnk}}$ each make contributions to $\overline{V^2_{eh}}$ and $\overline{V^2_f}$.

The final expression shows a much more complicated dependence of the fluctuations on the circuit and space-charge conditions than has hitherto been expected. The individual features of this dependency have been touched upon in the earlier Sections as they have been brought out by the equations, but, to recapitulate, it will be noticed that the space charge, when it controls the current, reduces the fluctuations. On the other hand, when it contributes an appreciable capacitive effect, it enhances the fluctuations and also gives rise to a flicker effect when the time-constant of the measuring amplifier is very small.

If we consider the effect of an inductance in parallel with the resistance and capacitance, the voltage $V(t)$ due to the passage of

one electron will become oscillatory, but the integral in eqn. (9) remains very nearly unchanged. Consequently we may expect the flicker effect to be observed at all mean frequencies, provided that the bandwidth is small enough. This aspect of the flicker effect has not previously been considered, but there is some experimental evidence, albeit by force of circumstances of only limited extent, that the flicker effect is in fact observed at all frequencies up to 20 kc/s when the bandwidth is sufficiently reduced.

The enhancement of the fluctuations and the flicker effect described here can be expected wherever a current sets up a space charge. It may therefore be possible to explain at least part of the flicker effect observed in semi-conductors—germanium diodes, for example—in a similar manner.

It is very likely that the crude assumption leading to eqn. (6) for the effect of n on C_e will give rise to far too large an estimate of the flicker effect. It has been used, however, so as not to cloud the physical picture with complicated mathematics. A better assumption would be to put

$$C_e = (\bar{n} - n)^2 k$$

which would give a much reduced estimate. However, the number $(\bar{n} - n)$ is in general fractional, and to use it would be inconsistent with the methods used in this paper.

Finally, it should be noticed that, although we have used Fig. 2 to indicate that a differential coefficient, in this case $\partial Q_s/\partial V$, fluctuates when noise voltages are under consideration, we have in fact used p.g.f.'s which are not dependent on the fluctuation voltage, δV. In other words, we have tacitly assumed that the fluctuation voltages occupy only a negligible range of abscissae. This is clearly not a satisfactory position in which to leave the matter. We can improve the accuracy of our considerations by making the parameters a and c of eqn. (17) functions of δV, such that the average value of the differential coefficient remains constant but its fluctuations decrease as δV increases. This opens up the possibility, not hitherto available, of letting δV include not only fluctuation voltages but also applied signals of varying magnitude, so enabling the masking of signals by noise to be studied more realistically.

(8) REFERENCES

(1) Bull, C. S.: 'The Capacitance between Planar Electrodes in the Presence of Space Charges', *Proceedings I.E.E.*, Paper No. 2333 R, July, 1957 (**104** B, p. 374).

(2) Bull, C. S.: 'The Noise of Valves', *British Journal of Applied Physics*, 1954, **5**, pp. 127 and 270.

(3) Campbell, N. R.: 'The Study of Discontinuous Phenomena', *Proceedings of the Cambridge Philosophical Society*, 1909, **15**, pp. 117 and 310.

(4) Whittaker, J. M.: 'The Shot Effect of Showers', *ibid.*, 1937, **33**, p. 451; 1938, **34**, p. 158.

(5) Rowland, E. N.: 'The Theory of the Mean Square Variations of a Function formed by Adding Known Functions', *ibid.*, 1936, **32**, p. 580.
'Note on Fluctuations and the Shot Effect', *ibid.*, 1938, **34**, p. 329.

(6) Rice, S. O.: 'The Mathematical Analysis of Random Noise', *Bell System Technical Journal*, 1944, **23**, p. 282.

621.317.73.029.64

The Institution of Electrical Engineers
Paper No. 2467 R
Mar. 1958

AN INSTRUMENT FOR THE MEASUREMENT OF SURFACE IMPEDANCE AT MICROWAVE FREQUENCIES

By A. E. KARBOWIAK, Ph.D., B.Sc.(Eng.), Associate Member.

(*The paper was first received 25th June, and in revised form 12th October*, 1957.)

SUMMARY

A theory is given of an instrument for the measurement of surface impedance at microwave frequencies.

Two prototype instruments—one for 6 Gc/s and one for 34 Gc/s—are described, and the measurement procedure is explained. Basically the surface impedance ($Z_s = R_s + jX_s$) is deduced from the resonant conditions of the cavity operated simultaneously in H_{01}- and E_{11}-modes. In particular, the cavity bandwidth is a measure of the surface resistance R_s, and the surface reactance, X_s, is deducible from the separation between the H_{01} and E_{11} responses.

As an illustration of the use of the instruments a number of experimental results are cited.

LIST OF PRINCIPAL SYMBOLS

l = Resonator length.
s = Radius of the resonator.
$k_0 = 2\pi/\lambda_0$ = Free-space wave number of a perfect cavity.
h_0 = Cut-off coefficient of a perfect cavity.
β_0 = Axial coefficient of a perfect cavity.
$\delta k, \delta h, \delta\beta, \delta f_H, \delta f_E$ = Various perturbation parameters.
$k = k_0 + \delta k$ = Wave number of the (imperfect) cavity.
$h = h_0 + \delta h$ = Cut-off coefficient of the cavity.
$\beta = \beta_0 + \delta\beta$ = Axial coefficient of the cavity.
f_0 = Resonant frequency of the perfect cavity.
f_E, f_H = Resonant frequency of the cavity in the E_{11}- and H_{01}-mode, respectively.
$b_E = 1/Q_E, b_H = 1/Q_H$ = Bandwidth of the cavity in the E_{11}- and H_{01}-modes, respectively.
$Z_s = R_s + jX_s$ = Surface impedance normalized with respect to Z_0 (= 377 ohms).
δ = Separation between the E_{11} and H_{01} resonant frequencies.

(1) INTRODUCTION

At microwave frequencies surface impedance is a particularly important physical property of a surface, in an analogous manner as at relatively low frequencies the impedance of a component in an electric circuit is an important physical property of the component.

In electric circuits, if the impedances of all the components and the manner in which the components are connected are known, the electric behaviour of the circuit as a whole can be computed with precision. Likewise, at any particular frequency the performance of a microwave circuit is (at least in principle) predictable from its dimensions, its mode of operation and the surface impedance of all the current-carrying surfaces. Thus, for example, the phase-change and the attenuation coefficients of a waveguide are calculable, at any particular frequency and for any mode of operation, from the measurements of the wave-guide cross-section and the surface impedance of its walls. In a similar manner the resonant frequency and the Q-factor of a microwave cavity are completely determined for any particular mode of operation by the cavity geometry and the surface impedance of its walls.

In general, surface impedance is not a constant but is a function of frequency and often also of many other factors; this is not unlike the impedance of a component in an electric circuit—for example, the impedance of a condenser is over a limited frequency band inversely proportional to frequency.

The necessity of some means of measuring the surface impedance is apparent, particularly since in most practical cases the analytical difficulties in computing the value of the surface impedance for a given set of its physical constants are considerable (e.g. rough surfaces).

There are a number of well-known methods for the determination of surface loss (essentially proportional to surface resistance, or the real part of surface impedance) to a reasonable degree of accuracy; the attenuation measurements on waveguides are just one illustration of such a method. The measurement of surface reactance (the imaginary part of surface impedance), however, is a much more difficult proposition, and there are no known methods of determining this quantity satisfactorily for practical needs.

In what follows, a description of a surface-impedance measuring instrument is given. Two instruments were designed, one for the 6 Gc/s band and the other for operation at a free-space wavelength of 8·7 mm, and the use of the instruments is illustrated by experimental results. Using these instruments it is possible to determine both the real and imaginary parts of the surface impedance of a cylindrical sample forming part of the cavity. In principle, the instrument consists of a cylindrical cavity excited in the H_{01}- and E_{11}-modes: the surface resistance is determinable from the Q-factor, and the surface reactance from the separation between the H_{01}- and E_{11}-responses.

(2) PRINCIPLE OF OPERATION OF THE INSTRUMENT

The instrument consists of a cylindrical cavity of circular cross-section, as shown in Fig. 1, where (*a*) and (*a'*) are the two

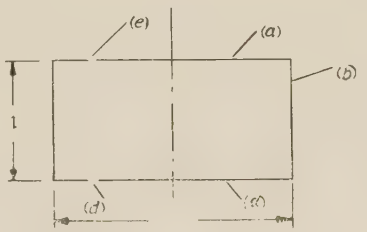


Fig. 1.—Schematic of the instrument.

removable end-plates and (*b*) is the curved wall of the cavity (hereafter referred to as the 'wall') which is the sample under test.

There are means for coupling the energy into (*c*) and out of the cavity (*d*), but the constructional details of the coupling are a

Written contributions on papers published without being read at meetings are invited for consideration with a view to publication.

Dr. Karbowiak is with Standard Telecommunication Laboratories, Ltd.

matter of engineering convenience in relation to the frequency of operation of the instrument.

The dimensions of the cavity (diameter $D = 2s$, length l) and the type of coupling are so chosen that the cavity can be operated at the design frequency in either the H_{01}- and E_{11}-modes, or both, without the interference from any other modes.

To explain the manner in which the instrument works, suppose that it is required to have an instrument working in the vicinity of a frequency f_0. For any such frequency the dimensions D and l can be so chosen that in its vicinity the cavity will only respond to H_{011}- and E_{111}-modes of oscillation.

If the cavity were perfect, the following relation would be true:[1]

$$k_0^2 = h_0^2 + \beta_0^2 \quad . \quad . \quad . \quad . \quad . \quad . \quad (1)$$

where

$$\left.\begin{aligned} k_0 &= \frac{2\pi}{\lambda_0} = \frac{2\pi f_0}{c} \\ \beta_0 &= \frac{\pi}{l} \\ h_0 &= \frac{3\cdot 832}{s} \end{aligned}\right\} \quad . \quad . \quad . \quad . \quad (2)$$

Eqn. (1) gives a relation between the dimensions of a cavity and the resonant frequencies of the H_{011}- and E_{111}-modes, which in the case of the above perfect cavity are identical. Further, the Q-factor ($Q = 1/b$, where b is the half-power bandwidth) of such a cavity is infinite.

The cavity is termed 'perfect' because the surface impedance of all its walls [(a), (a') and (b)] have been assumed zero.

Physical cavities are referred to as 'imperfect' because of their non-vanishing value of the surface impedance. If the surface impedance is small, which is true of all practical cavities, eqn. (1) is replaced by[1]

$$k^2 = h^2 + \beta^2 \quad . \quad . \quad . \quad . \quad . \quad . \quad (3)$$

where

$$\left.\begin{aligned} k &= k_0 + \delta k \\ h &= h_0 + \delta h \\ \beta &= \beta_0 + \delta\beta \end{aligned}\right\} \quad . \quad . \quad . \quad . \quad . \quad (4)$$

To consider a homogeneous physical cavity we can imagine the cavity walls to have a surface impedance $Z_s = R_s + jX_s$, and the following formulae can be shown to be true.

Let

$f_H = f_0 + \delta f_H$

= Resonant frequency of the cavity in the H_{01}-mode

$Q_H = \dfrac{1}{b_H}$

= Q-factor of cavity in the H_{01}-mode

b_H = Half-power bandwidth of the cavity in the H_{011}-mode and f_E, Q_E, b_E be similar quantities pertaining to the resonance of the cavity in the E_{11}-mode. It can then be shown that[1]

$$\left.\begin{aligned} \frac{\delta f_E}{f_0} &= -\left(\frac{1}{s} + \frac{2}{l}\right)\frac{X_s}{k_0} \\ \tfrac{1}{2}b_E &= \left(\frac{1}{s} + \frac{2}{l}\right)\frac{R_s}{k_0} \end{aligned}\right\} \quad . \quad . \quad . \quad . \quad . \quad (5)$$

for the E_{11}-mode, and

$$\left.\begin{aligned} \frac{\delta f_H}{f_0} &= -\left[\frac{1}{s}\left(\frac{h_0}{k_0}\right)^2 + \frac{2}{l}\left(\frac{\beta_0}{k_0}\right)^2\right]\frac{X_s}{k_0} \\ \tfrac{1}{2}b_H &= \left[\frac{1}{s}\left(\frac{h_0}{k_0}\right)^2 + \frac{2}{l}\left(\frac{\beta_0}{k_0}\right)^2\right]\frac{R_s}{k_0} \end{aligned}\right\} \quad . \quad . \quad . \quad (6)$$

for the H_{01}-mode.

Evidently the bandwidth of either E_{11}- or H_{01}-response is a measure of R_s and presents no difficulty. The quantity X_s would, however, be deducible from the resonant frequency of the cavity in the E_{11}- or H_{01}-mode if the frequency, f_0, of the perfect cavity were known.

The frequency f_0 cannot, however, be measured, and the best that can be done is to deduce it from the measurements of the cavity using eqns. (2). Notwithstanding, the small errors in the measurement of $s(\delta s)$ and $l(\delta l)$ lead to an error in the deduced resonant frequency f_0 given by

$$-\frac{\delta f_0}{f_0} = \left(\frac{h_0}{k_0}\right)^2\frac{\delta s}{s} + \left(\frac{\beta_0}{k_0}\right)^2\frac{\delta l}{l} \quad . \quad . \quad . \quad . \quad (7)$$

Suppose that Z_s is the surface impedance of a plain copper surface at 9 Gc/s (6×10^{-5}). It follows from eqns. (6) and (7) that to be able to measure it to a sensible degree of accuracy the quantities δs and δl would have to be known to within a few millionths of an inch, a prohibitive requirement.

Here it will be convenient to introduce an additional quantity δ, defined by

$$\begin{aligned} \delta = \frac{\delta f}{f_0} &= \frac{\delta f_H - \delta f_E}{f_0} \\ &= \text{Separation between the E and H resonances} \\ &= \left[\frac{1}{s}\left(\frac{\beta_0}{k_0}\right)^2 + \frac{2}{l}\left(\frac{h_0}{k_0}\right)^2\right]\frac{X_s}{k_0} \quad . \quad . \quad . \quad . \quad . \quad . \quad (8) \end{aligned}$$

It will be observed that this is independent of δs and δl and is a measure of the required quantity X_s.

Fundamentally, therefore, in the case of the surface-impedance measuring instrument the surface reactance is deduced from the difference between the resonant frequencies of the cavity when excited in the H_{01}- and E_{11}-modes, and this is obtainable with great precision from the experimental observations.

Suppose that the instrument is made of a homogeneous material, so that all its walls have equal surface impedances Z_1. To determine its value only two measurements are needed: b_H, the half-power bandwidth of the H_{01}-response, giving

$$R_1 = \frac{k_0}{2}\left[\left(\frac{h_0}{k_0}\right)^2\frac{1}{s} + \left(\frac{\beta_0}{k_0}\right)^2\frac{2}{l}\right]^{-1} b_{H1} \quad . \quad . \quad . \quad (9)$$

and δ, the separation between the H_{01}- and E_{11}-responses, giving

$$X_1 = k_0\left[\left(\frac{\beta_0}{k_0}\right)^2\frac{1}{s} + \left(\frac{h_0}{k_0}\right)^2\frac{2}{l}\right]^{-1}\delta_1 \quad . \quad . \quad . \quad (10)$$

Evidently, both R_1 and X_1 can be determined to a high degree of accuracy.

Usually it is desirable to measure a third quantity: b_E, the half-power bandwidth of the E_{11}-response, giving

$$R_1 = \frac{k_0}{2}\left(\frac{1}{s} + \frac{2}{l}\right)^{-1} b_E \quad . \quad . \quad . \quad . \quad . \quad (11)$$

Eqn. (11) should lead, within experimental errors, to the same answer as eqn. (9), unless the surface impedance is aniso-

ipic or the cavity defective in some way. In this manner a ck on these factors can be kept.

Having determined the behaviour of the cavity with all its ls of surface impedance Z_1, the sample can now be replaced one to be measured (of surface impedance Z_2) and the asurement repeated.

Let the relevant quantities be denoted by

b_{H2} = Bandwidth in the H_{01}-response

b_{E2} = Bandwidth in the E_{11}-response

δ_2 = Frequency separation between H_{01}- and E_{11}-responses

en, using eqns. (9), (10) and (11), the following expressions the unknown impedance $Z_2 = R_2 + jX_2$ can be derived:

$$\left.\begin{aligned} R_2 &= R_1 + sk_0(b_{H2} - b_{H1})\left(\frac{k_0}{h_0}\right)^2 \\ X_2 &= X_1 + 2sk_0(\delta_2 - \delta_1)\left(\frac{k_0}{\beta_0}\right)^2 \end{aligned}\right\} \quad . \quad . \quad (12)$$

lh an additional expression for R_2 given by

$$R_2 = R_1 + sk_0(b_{E2} - b_{E1}) \quad . \quad . \quad . \quad . \quad (13)$$

Hence the value of the unknown surface impedance can be iced.

t is only right to stress that the measurement is an absolute e and that the instrument is not a surface-impedance com- rator. The quantity Z_1, which can be obtained from an ependent measurement as described above, is simply a con- nt of the instrument.

A typical response curve of the instrument, with all the relevant antities indicated is shown in Fig. 2.

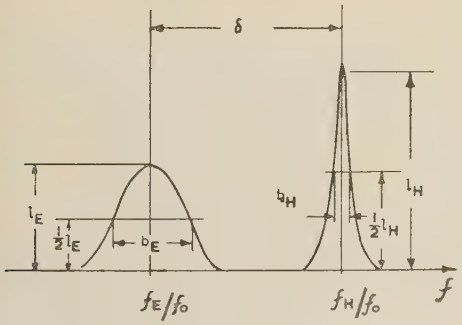


Fig. 2.—Resonant response of the instrument.

(3) DESCRIPTION OF THE INSTRUMENTS

(3.1) General

The underlying principles and the *modus operandi* of the trument have been explained in Section 2, but the instrument its practical form must be improved and elaborated in several pects if a versatile defect-free and reliable instrument is to ult. One of such aspects is the contact impedance between sample and the cavity end-plates.

The currents in the walls of the cavity, and associated with H_{01}-mode, flow in circles concentric with the axis of the vity; consequently there is no current flowing across the ction between the sample and the rest of the cavity. The onant conditions of the cavity are therefore independent of the lar y of this junction.

With the E_{11}-mode, however, the situation is different. Here currents are flowing across the junction, and consequently the conditions of resonance are to a noticeable extent influenced by its nature, which for all practical cases is largely unpredictable.

In effect, as far as the performance of the instrument is concerned, the resonant frequency and Q-factor of the E_{11}-mode are substantially modified by the relevant contact impedance but the resonance conditions of the H_{01}-mode remain unaffected. This phenomenon manifests itself, for example, in the lack of agreement between the results for R_1, as obtained from measurements of b_H and b_E [eqns. (9) and (11)]. The consequential experimental error in determining the surface impedance Z_2 is often objectionable.

To bridge the difficulty the cavity is constructed of three parts made in definite proportion, as shown in Fig. 3. The total

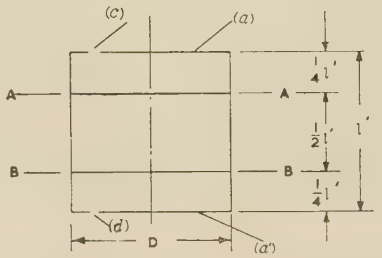


Fig. 3.—Improved version of the instrument.

length of the cavity is twice that shown in Fig. 1 and the cavity is operated in the H_{012}- and E_{112}-modes. The top part of the cavity, above the plane AA, is made of one piece of metal and so is the part of the cavity below the plane BB. The part of the cavity between the planes AA and BB is the sample under investigation. In this construction of the instrument there are no currents flowing across any mechanical junctions, and consequently any errors associated with contact impedances are eliminated.

In this improved form of construction, if the upper and lower parts of the cavity each have a surface impedance Z_1 and the middle part (the sample) has a surface impedance Z_2, eqns. (9)–(11) are modified to read:

$$\left.\begin{aligned} \tfrac{1}{2}b_H &= \tfrac{1}{2}\left(\frac{h_0}{k_0}\right)^2 \frac{R_2 + R_1}{sk_0} + 2\left(\frac{\beta_0}{k_0}\right)^2 \frac{R_1}{lk_0} \\ \tfrac{1}{2}b_E &= \tfrac{1}{2}\frac{R_2 + R_1}{sk_0} + 2\frac{R_1}{lk_0} \\ \delta &= \tfrac{1}{2}\left(\frac{\beta_0}{k_0}\right)^2 \frac{X_2 + X_1}{sk_0} + 2\left(\frac{h_0}{k_0}\right)^2 \frac{X_1}{lk_0} \end{aligned}\right\} \quad . \quad (14)$$

where l is the total length of the cavity.

Clearly, if Z_1 is known from an independent measurement, Z_2 can be determined from the above equations.

It is evident that when the surface impedance is such that the reactive component is of the order of the resistive component or less, the two responses (one due to the H_{01}-mode and the other due to the E_{11}-mode) merge into each other (Fig. 4). Conse-

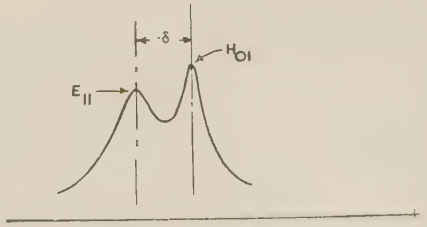


Fig. 4.—Typical response of the instrument.

quently, it becomes more and more difficult to determine with accuracy their positions and Q-factors. To improve the measurement accuracy under such conditions the cavity is modified in a number of ways.

The coupling means (*c*) and (*d*) are altered so that the proportion of the H_{01}-mode can be varied relative to the E_{11}-mode. At frequencies not too high (say below 10 Gc/s) this is achieved by feeding the cavity through a coaxial cable and a loop-probe, as shown in Fig. 5. In addition, the whole probe assembly is

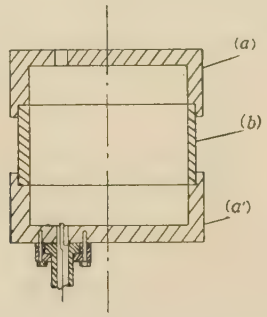


Fig. 5.—Schematic of the 6 Gc/s instrument.

rotatable in its mounting and thus the plane of the polarization of the probe can be varied. Clearly, if the probe is rotated so that the plane of the loop-probe also contains the resonator axis, only the E_{11}-mode is launched; in any other position of the probe the proportion of the H_{01}-mode is determined by the angle between the plane of the probe PP and the plane XX, as shown in Fig. 6. Analogous arguments apply to the detector probe.

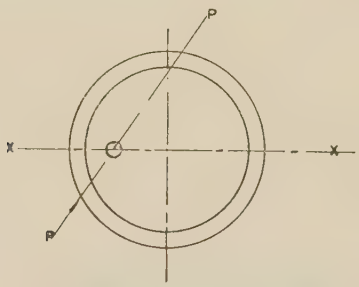


Fig. 6.—Geometrical position of the probe relative to the cavity.

At frequencies higher than about 10 Gc/s it may be more convenient to have, instead of coaxial, waveguide coupling means. In such a case the construction shown in Fig. 7 is preferable. Here, rectangular waveguide feeds energy via a conventional rectangular H_{01}- to circular H_{11}-mode transducer (*e*) to the coupling hole (*d*). The transducer is rotatable in its mounting (*f*) and is kept in position by a retaining ring (*g*).

If the transducer (*e*) is rotated, so that the plane PP of the feeding waveguide contains the axis of the cavity, only the E_{11}-mode will be launched; in any other position the proportion of the H_{01}-mode will be determined by the angle between the planes PP and XX (Fig. 6). Similar remarks apply to the detector probe (*c*).

Furthermore, the two parts of the instrument (*a*) and (*a'*) can be rotated relative to each other, which greatly facilitates the separation of the E- and H-responses for the following reason:

The intensity of the H_{01}-mode is independent of the angular position of the detector probe relative to the feed probe, but the intensity of the E_{11}-mode varies sinusoidally with the angle

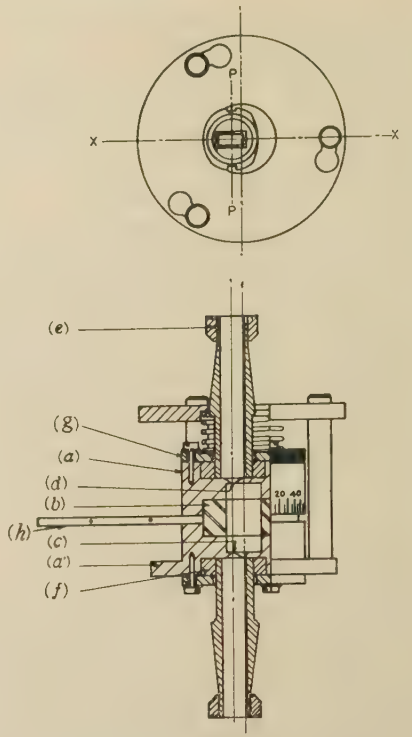


Fig. 7.—The 8·7 mm instrument.
(*a*), (*e*) End-caps.
(*b*) Sample.
(*c*), (*d*) Coupling aperture.
(*e*) Output waveguide.
(*f*) Input waveguide.
(*g*) Upper retaining ring.
(*h*) Sample holder.

between the probes; consequently two positions (diametrical opposite) can be found in which there is no coupling betwe the feed and the detector via the E_{11}-mode.

Using this refined form of the instrument, the two respons one due to the E_{11}-mode and the other due to the H_{01}-mod can be separated easily, even for very small values of surfa impedance, by a procedure described below.

(3.2) An Instrument designed for the 35 Gc/s Band

The instrument shown in Figs. 7 and 8 is suitable for measur ments in the 35 Gc/s band. The diameter of the cavity is ju

Fig. 8.—Exploded view of the 8·7 mm instrument.

under 0·5 in and the length of the cavity (*l*) is a little ov 0·75 in. These dimensions were chosen, with the help of mode chart, so that in the frequency range about 500 Mc above and below the resonant frequency of the H_{01}-mo (34·5 Gc/s) the cavity is free from any other modes of oscillatio except the E_{11}-mode. The margin of 500 Mc/s is sufficient

e instrument is intended to measure surface impedance whose eactive part does not exceed about 10^{-2} (i.e. approximately ohms), which is considered ample for the purposes intended. owever, the instrument could be designed, if required, for a rger frequency margin.

The instrument is fed from a swept-frequency microwave nerator of a suitable frequency through the lower transducer d coupling hole (Fig. 7), while a crystal detector is connected the end of the upper transducer. The rectified crystal output fed through an amplifier to an oscilloscope whose time-base is nchronized with the voltage responsible for frequency sweeping the microwave generator.

A typical response curve for a moderate value of surface actance is shown in Fig. 10.

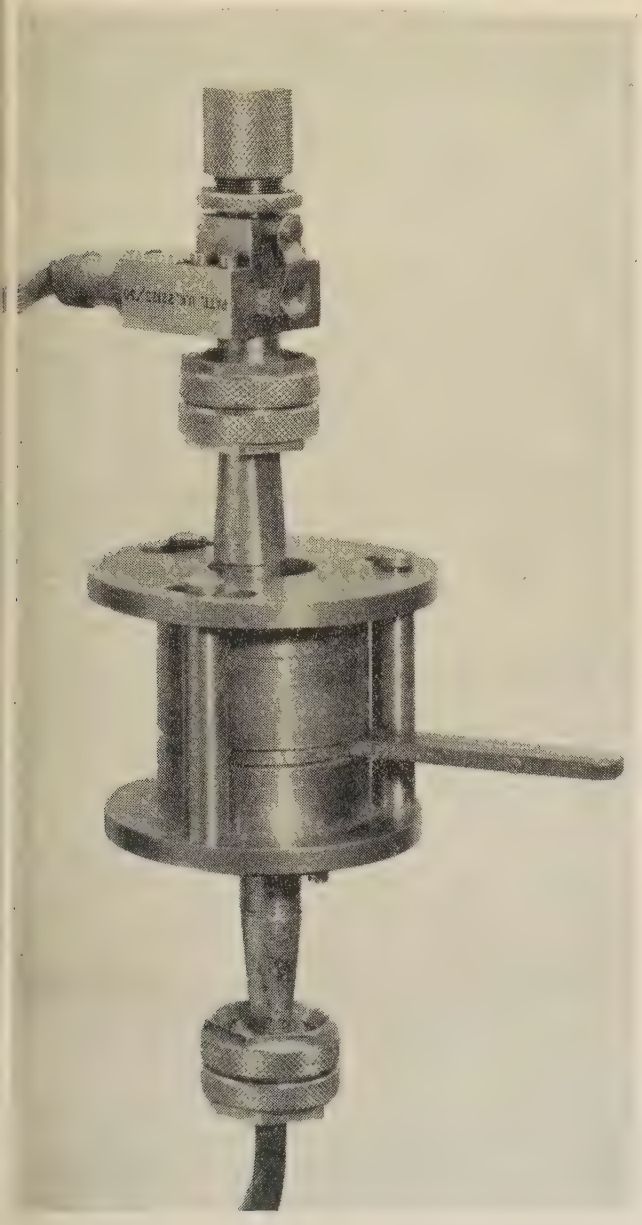


Fig. 9.—The 8·7 mm instrument assembled for test.

The experimental procedure is as follows. The instrument, ich is preferably made of coin-silver, is fitted with a coin-silver mple and connected up as shown in Fig. 9. With the upper ate turned with respect to the lower one, so that the coupling means (c) is immediately above the coupling means (d), the instrument is brought into resonance, whereupon a display similar to that shown in Fig. 10 will be obtained. The feeds

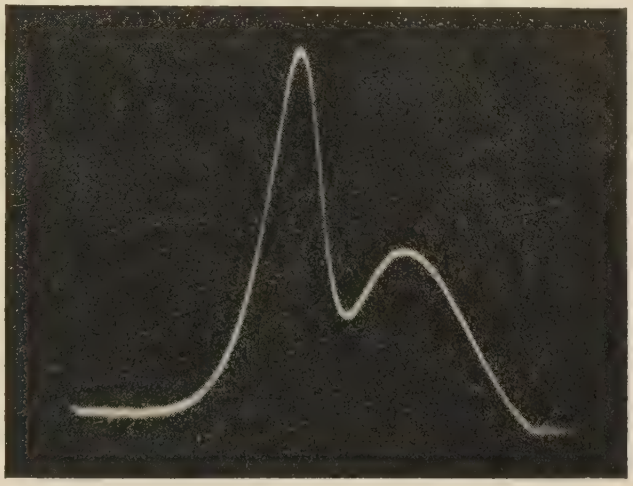


Fig. 10.—Typical resonance curve of the 8·7 mm instrument showing a separation between the H_{01} and E_{11} responses of approximately 7 Mc/s.

are then turned in their respective mountings until the response due to the H_{01}-mode disappears, and the measurement of the bandwidth [giving the quantity R_1 through eqn. (13)] and the resonant frequency are carried out.

Subsequently the upper part of the cavity is turned round (by 90°) until the response due to the E_{11}-mode disappears. The probes are then adjusted in their respective mountings until the response (due to the H_{01}-mode) is at its maximum, and the measurement of the resonant frequency and Q-factor [giving the quantity R_1 through eqn. (9)] are carried out.

To measure the surface impedance of an unknown sample the master sample is replaced by that to be measured. The procedure is as described above in connection with the master sample, and the measured values of b_E, b_H and δf are used in conjunction with eqns. (9), (10), (11) and (14) to determine the real and imaginary parts of the surface impedance Z_2.

The samples are provided with a grip-on sample holder (h) by means of which the sample can be turned round to various positions, so that a mean of several measurements may be obtained should the sample be irregular.

Clearly, great care must be exercised in the making of samples, particularly where a high degree of accuracy of measurement is required. For example, an elliptical sample leads to a split in the E_{11} resonance line in the case of substantial ellipticity, but only to an apparent broadening of the resonant response in the case of small ellipticity. An irregular sample will usually exhibit a lower Q-factor than a symmetrical one, and, furthermore, in the case of irregular samples the quantity δf becomes a function of the angle between the feed and detector probes.

(3.3) An Instrument for Operation at about 6 Gc/s

The instrument for the 6 Gc/s band is of a construction somewhat different from the one described above. In principle it could have been a scaled version of that instrument, but, because at 6 Gc/s waveguides become too bulky, coaxial feeds are preferable. In consequence the probes take the shape of little loops, as indicated schematically in Fig. 5.

A typical response curve of the instrument would be similar to that shown in Fig. 10.

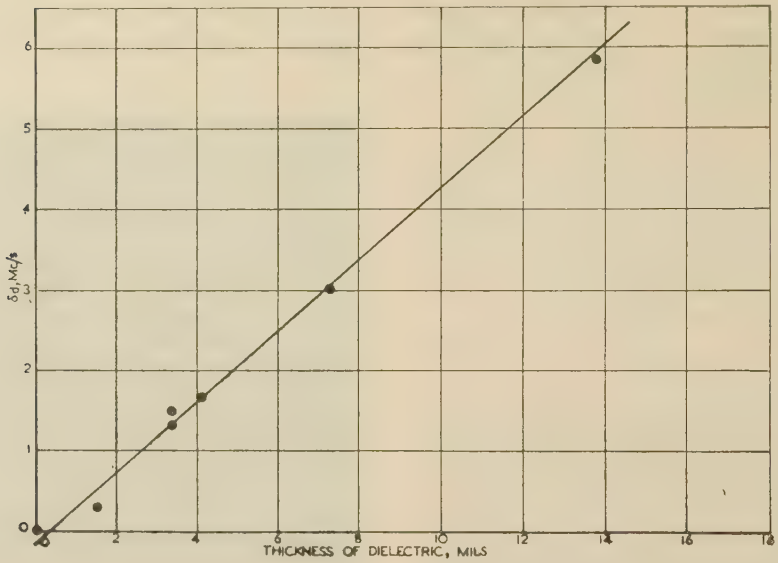


Fig. 11.—δ_d as a function of the thickness, t, of the dielectric.
Resonant frequency in the H_{01}-mode approximately 6·3 Gc/s.

The measurement procedure is analogous to that described in connection with the 35 Gc/s surface-impedance measuring instrument.

(4) DISCUSSION OF THE USE OF THE INSTRUMENTS AND SOME EXPERIMENTAL RESULTS

(4.1) An Experiment using the 6 Gc/s Instrument

For a certain microwave application a detailed study of the effect on surface impedance of thin skins of dielectrics over a copper surface was required.

For this purpose the 6 Gc/s instrument was employed and a number of cylindrical samples were made of copper, with the required dielectric deposited thinly (by the process under investigation) on the inner surfaces. In this way various dielectrics and methods of application were put to test, and the material and method were chosen to meet the required specifications, in a minimum amount of time.

Theoretical investigations[1,2] show that a thin skin of dielectric over a metallic surface has a negligible surface impedance as far as H-waves are concerned, but a significant surface impedance for E-waves. Referred to the surface of the metal backing the dielectric layer, it is given by[1]

$$Z_d = jk_0 t\left[\left(\frac{h_1}{k_1}\right)^2 - \left(\frac{h_0}{k_0}\right)^2\right] \quad . \quad . \quad . \quad . \quad (15)$$

which can be transformed into

$$Z_d = R_d + jX_d = jt\beta_0\left(\frac{\beta_0}{k_0}\right)\left(1 - \frac{1}{\epsilon_r}\right) \quad . \quad . \quad . \quad (16)$$

If we put $\epsilon_r = \epsilon_1(1 - j\tan\delta)$, then

$$\left.\begin{aligned} R_d &= t\beta_0\left(\frac{\beta_0}{k_0}\right)\frac{\tan\delta}{\epsilon_1} \\ X_d &= t\beta_0\left(\frac{\beta_0}{k_0}\right)\left(1 - \frac{1}{\epsilon_1}\right) \end{aligned}\right\} \quad . \quad . \quad . \quad . \quad (17)$$

The total surface impedance, however, is larger by the amount of the surface impedance of the metal surface, given by[1]

$$Z_m Z_0 = (1 + j)\sqrt{\left(\frac{\pi f\mu_m}{\sigma}\right)} \quad . \quad . \quad . \quad . \quad (18)$$

This, however, can be neglected only if the increase in bandwidth of the cavity and the increase, δ_d, in the separation between the E- and H-modes are noted.

Evidently, the graph of δ_d versus the thickness of the dielectric layer should show a straight line passing through the origin, its slope being a measure of the surface reactance of the dielectric coating.

The experimental results for this case are shown in Fig. 11. The dielectric skin was made of polystyrene deposited from a chemical solution, and the dielectric constant calculated from the slope of the experimental line is 2·6. Throughout, within experimental errors, Q_H was unaffected, while Q_E was a decreasing function of the thickness of the dielectric coating; Fig. 12 shows

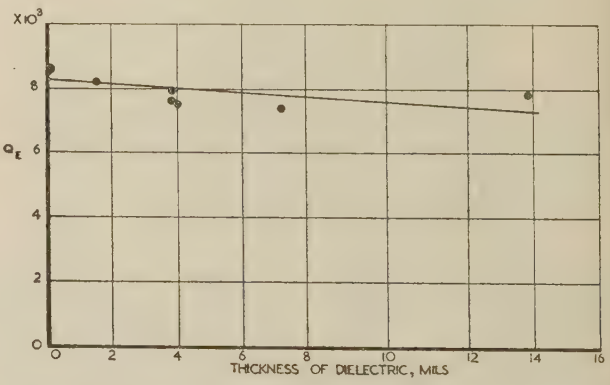


Fig. 12.—Q_E as a function of dielectric coating thickness.

the experimental results. The experimental points, it will be noted, are rather erratic and the loss tangent deduced from the results is of the order of 0·01, which is considerably greater than that for solid polystyrene (0·0007): this is attributable to contamination of the dielectric by the solvent.

(4.2) An Experiment in the 35 Gc/s Band

In connection with some work carried out on waveguides for long-distance transmission,[5] it was observed that the performance of a long length of circular waveguide operated in the H_{01}-mode

largely affected by the surface impedance of the waveguide all; and with a view to improving the performance of the aveguide when subject to bends it was thought desirable to rrugate its surface.

In order to achieve the necessary performance with a moderate ort and without unduly increasing the attenuation of the aveguide, a knowledge of the surface impedance of rfaces (with a triangular profile) as a function of surface rameters was required. To that end a number of suitable all cylindrical samples were prepared.

All samples were turned individually, in a lathe, to the required easurements, and the corrugations were formed by cutting a read (using a suitable ground tool) on the inner surface of e sample. Subsequently each sample was measured in the Gc/s instrument.

As a matter of interest and for the purpose of subsequent scussion, apart from b_H (bandwidth at resonance in the H_{01}-ode), b_E (bandwidth at resonance, in the E_{11}-mode) and δ eparation between the E_{11}- and H_{01}-responses), two more antities were measured:* δ_H (change in the H_{01} resonant fre-ency when a threaded sample is substituted for the unthreaded) d δ_E (corresponding change in the E_{11} resonant frequency). e experimental results are shown in Figs. 13 and 14.

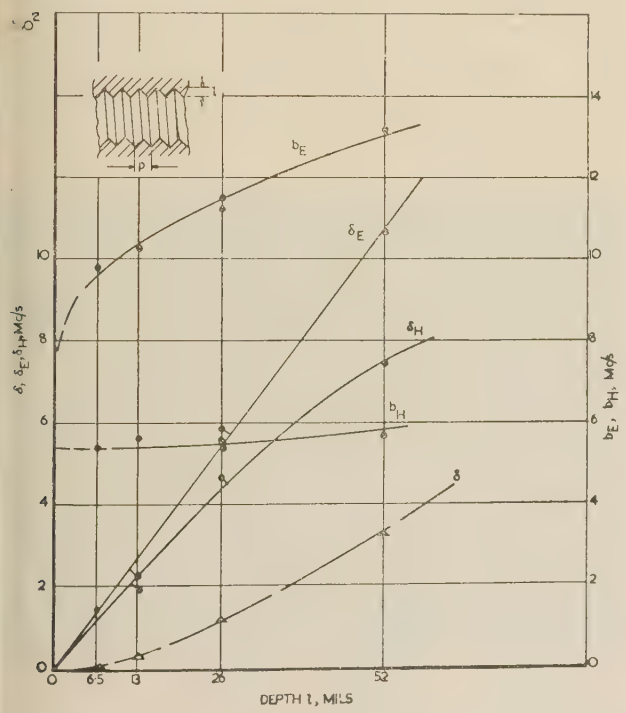


Fig. 13.—b_E, b_H, δ_E, δ_H, δ plotted against depth of corrugation, l, for a constant pitch of 12 turns per inch.
$p = 0{\cdot}083$ in, $\lambda_0 = 8{\cdot}7$ mm.

Fig. 13 shows the variations of δ_H, δ_E, b_H, b_E and δ as functions f the depth of corrugations for a fixed pitch (0·083 in), and ig. 14 shows the variations of the same quantities with pitch for constant depth of corrugations (0·013 in).†

Before attempting to interpret the experimental results it is perative to observe that the surfaces investigated are aniso-

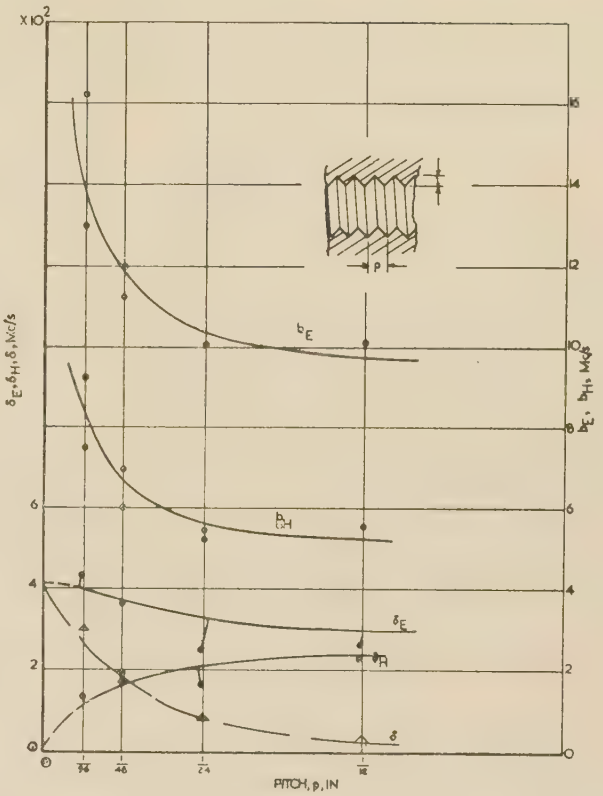


Fig. 14.—b_E, b_H, δ_E, δ_H, δ plotted against pitch, p, for a constant depth of corrugations, l.
$l = 13$ mils, $\lambda_0 = 8{\cdot}7$ mm.

tropic[3] and also that the principal axes of anisotropy do not coincide with the co-ordinate axes of the cavity. Since, however, the angle of anisotropy enters into the computations as the square of the tangent of the angle[3] and the angles involved are small, it is permissible to neglect its effect on the experimental results. However, it is essential to consider the anisotropic nature of the surface.

Using the formulae given above and some formulae quoted elsewhere,[1] it can be shown that, if we denote the anisotropic surface impedance $[Z']$ by its two components Z_c' (circumferential) and Z_a' (axial), and let δ_E be the difference between the resonant frequency of the instrument with a threaded sample and that with a smooth sample, and δ_H a similar quantity pertaining to the resonance in the H_{01}-mode, while Δb_E and Δb_H are the increases in the bandwidth of the cavity under the above conditions for the E_{11}- and H_{01}-modes, respectively, the following expressions can be derived:

$$\left.\begin{aligned} \delta_E &= \frac{1}{2}\frac{X_a}{sk_0} \\ \Delta b_E &= \frac{R_a}{sk_0} \\ \delta_H &= \frac{1}{2}\left(\frac{h_0}{k_0}\right)^2 \frac{X_c}{sk_0} \\ \Delta b_H &= \left(\frac{h_0}{k_0}\right)^2 \frac{R_c}{sk_0} \end{aligned}\right\} \quad . \quad . \quad . \quad . \quad . \quad (19)$$

where
$$[Z'] = [Z] + Z_m[1] \quad . \quad . \quad . \quad . \quad . \quad (20)$$

* It will be appreciated that these quantities cannot be measured to anything like e accuracy of measurement of δ, for reasons given in Section 2, but in the present st nce it is nevertheless instructive to measure them. Since the surface reactances vo ved are rather large the percentage errors are not too excessive.

† Similar experiments were carried out on samples of different proportions, but the su s will not be recorded here.

i.e. $[Z]$ = Increase in surface impedance of the anisotropic surface above the value Z_m, the surface impedance of the unthreaded surface.

The increase in the separation between the E_{11}- and H_{01}-responses, δ, is now given by

$$\delta = \frac{1}{2}\left(\frac{\beta_0}{k_0}\right)^2 \frac{X_t}{sk_0} \quad . \quad . \quad . \quad . \quad . \quad . \quad (21)$$

where*

$$X_t = \left[\left(\frac{k_0}{\beta_0}\right)^2 X_a - \left(\frac{h_0}{\beta_0}\right)^2 X_c\right] \quad . \quad . \quad . \quad (22)$$

It is important to note that, whereas in the isotropic case δ is directly proportional to the surface reactance, the relation between δ and the principal anisotropic components, X_a and X_c, is not a simple one, but is given by eqns. (21) and (22).

Although the quantitative interpretation of the experimental results shown in Figs. 13 and 14 is not an easy task, the qualitative interpretation is quite simple, as follows:

With corrugated surfaces of a rather coarse pitch, the loss in the surface depends largely on whether the corrugations run across (E_{11}-wave) or along (H_{01}-wave) the lines of current flow. For this reason we would expect b_E to increase more rapidly than b_H; this phenomenon is exemplified in Fig. 13. With samples of fine pitch the results for b_E and b_H are rather erratic, owing to varying amounts of burr (difficult to avoid) formed during the thread-cutting process.

It is significant to note that for a given depth of corrugation b_H has an asymptotic value just in excess of 5 Mc/s (when the pitch becomes large), and this is the bandwidth of an uncorrugated cavity. On the other hand, b_E has an asymptotic value between 9 and 10 Mc/s, which is significantly higher than that of an uncorrugated cavity. The reason is again the surface finish, which, in the axial direction (E_{11}-wave), is significantly poorer than for a plain sample.

δ_E and δ_H show a large scatter in experimental results mainly because, as explained above, these quantities depend on surface reactance as well as on the effective diameter of the sample; the mean diameter is, however, difficult to assess with any accuracy and consequently the errors are large. The separation between the E- and H-responses, δ, on the other hand, depends solely on surface reactance and hence is determinable with accuracy. These results are thus in experimental confirmation of the discussion given in Section 2.

The interpretation of the shape of the δ_E and δ_H curves is somewhat complicated because these quantities, as mentioned above, depend on two different causes: the surface reactance and the effective diameter of the corrugated sample. Thus, for example, the curve of δ_E in Fig. 14 has two asymptotic values (*a*) for large values of pitch in relation to depth of corrugation when the effective radius of the corrugated sample is equal to the mean radius, and the surface reactance is substantially nil; and (*b*) for very small values of pitch, when the effective radius of the sample is equal to the radius of an imaginary cylinder which touches the high points of the corrugations, and the surface reactance is given substantially by $k_0 l_m$, where l_m is the mean depth of corrugation.

Between the two extremes the apportioning of δ_E between the two effects might not, at first, seem to be an easy matter. But a little more scrutiny leads to the following conclusion. To the first order of quantities (and this is our only concern), the circumferential component of the surface reactance, the only one affecting the H_{01}-mode, is substantially zero ($X_c \simeq 0$), and consequently δ_H arises from the change in the effective radius of the sample. This quantity, therefore, can be subtracted from δ_E

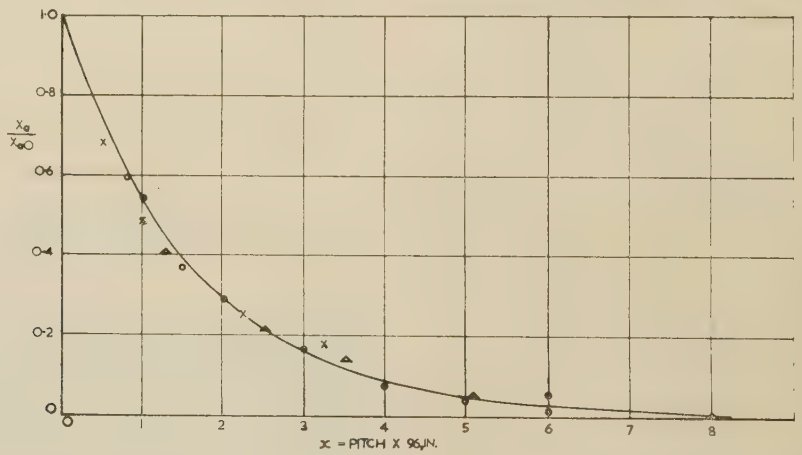


Fig. 15.—Normalized surface reactance of a corrugated surface with triangular corrugations.

$X_a/X_{a0} = \exp(-3p/\sqrt{l})$

- ■ Experimental results for $l = 6{\cdot}5$ mils.
- o Experimental results for $l = 13$ mils.
- △ Experimental results for $l = 26$ mils.
- × Experimental results for $l = 52$ mils.

and hence X_a can be determined. The relation between δ and X_a follows from eqns. (21) and (22):

$$X_a = 2\delta s k_0 \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (23)$$

The quantity to be measured is δ, the separation between the E_{11}- and H_{01}-responses, which, as discussed above, is determinable with accuracy.

In all, about 20 different samples were measured and the results shown in Figs. 13 and 14 are merely representative. On superficial examination the results seem to have little in common, but a detailed and careful examination shows that there is a definite correlation between the surface reactance of a corrugated surface (with a triangular profile) and the corrugation measurements. In fact, when suitably normalized, the averaged experimental points can be shown to fall on a common curve, as illustrated in Fig. 15.

If the pitch p and the depth of corrugations l are measured in inches, the experimental law can be shown to be

$$X_a = X_{a0} \exp(-3p/\sqrt{l}) \quad . \quad . \quad . \quad . \quad (24)$$

* It is to be noted that if $X_a = X_c$ (isotropic surface), $X_a = X_c = X_t$.

ere $X_{a0} = \frac{1}{2}k_0 l$ . . . . . (25)

= Surface reactance of the surface of very small pitch ($p \to 0$).

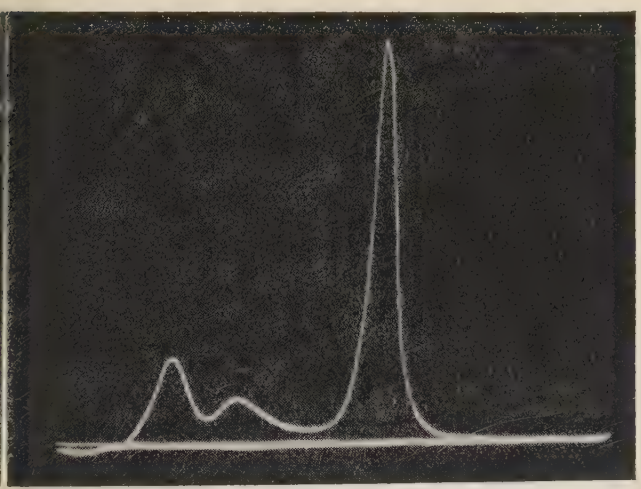
(4.3) Sample Uniformity

In Section 3.2 the need for uniform samples is stressed and the m 'uniformity' has been applied without discrimination to the ometry as well as to the surface-impedance function. In fact, small geometrical irregularity has the same effect on the onant conditions of a cavity as a suitable small irregularity in surface-impedance function. It transpires, further, that a ghtly non-uniform cavity can be treated as a heterogenous vity[4] and as such has several pecularities.

In the present application it is the circumferential heterogeneity at is of primary importance. For in such cavities the E_0- and $_0$-modes can be shown to be stable and the effective surface pedance is equal the mean value of the surface impedance nction; all other modes are, in general, unstable.

Consequently, with this instrument, the resonant conditions of e H_0-mode are determined by the mean value of the surface-pedance function. On the other hand, the E_{11}-mode is stable, unless the first Fourier harmonic of the surface-pedance function is absent. The instability manifests itself the split of the resonant line into the $_cE_{11}$- and $_sE_{11}$-modes[4] th a tight coupling between them.

The phenomenon has actually been observed on a number of casions, particularly with the samples fabricated from sheet-pper with an axial seam. A typical response of a cavity with ch a sample is shown in Fig. 16. The cavity evidently possesses



g. 16.—Response of the 6 Gc/s instrument provided with a sample having a defective axial seam.

high-Q H_{01}-response and a double E_{11}-response, as described ove. The reason is that unless the axial seam is made with treme care it leads to a substantial first-harmonic component the surface impedance function.

Clearly, the phenomenon of the split in the E_{11}-response into $_cE_{11}$ and $_sE_{11}$ could be put to advantage, e.g. to measure the rmonic content of a heterogeneous surface. In particular, in e last example the separation and the relative magnitudes of the $_cE_{11}$- and $_sE_{11}$-responses are a measure of the first-harmonic ntent of the surface-impedance function involved.[4]

(4.4) The Effect of Coupling Means on the Measurement Accuracy

Evidently the input and output couplings load the cavity and, at least in principle, modify the performance of the instrument in two respects: (*a*) the Q-factor will be lowered due to the loss in external circuits, and (*b*) the resonant frequency will be slightly altered due to the reaction of the external circuits (via the couplings) on the field inside the cavity.

There are two ways of eliminating this source of error: (*a*) by replacing the impedance of the end-plates by a fictitious value slightly modified to take account of the coupling loading, and using this value as the constant of the instrument (obtained experimentally), and (*b*) by simply decreasing the coupling constructionally below the significant level and increasing the gain of the receiver accordingly.

Although badly designed couplings could, undoubtedly, lead to significant errors in measurements, an experienced observer would without hesitation discover them, merely by rotating the probe in its mounting and observing whether the width of the resonant line or its position is changing by any significant amount.

(5) CONCLUSIONS

To measure the physical quantity surface impedance, two prototype instruments have been designed, developed and used for numerous measurements. To illustrate the use of the instrument a few experimental results have been quoted and the measurement procedure described.

Using the instruments the surface impedance can be determined with an accuracy of 5–20%, depending on conditions, but there is no reason why the accuracy should not be improved substantially. Thus, with some refinements in the measurement technique and using a more precision-built instrument it is hoped to realize a 5% (or better) measurement accuracy over the entire range of the instrument.

(6) ACKNOWLEDGMENTS

The author is indebted to a number of his colleagues in the laboratories, and in particular to Mr. L. Lewin, for helpful discussions, and Messrs. W. Pryciak and R. Skedd, for carrying out most of the experimental work.

Acknowledgment is made to Standard Telecommunication Laboratories Ltd. for facilities granted in the preparation of the manuscript and for permission to publish the paper.

(7) REFERENCES

(1) KARBOWIAK, A. E.: 'Theory of Imperfect Waveguides; the Effect of Wall Impedance', *Proceedings I.E.E.*, Paper No. 1841 R, September, 1955 (**102** B, p. 698).

(2) KARBOWIAK, A. E.: 'Theory of Composite Guides, Stratified Guides for Surface Waves', *ibid.*, Paper No. 1659 R, July, 1954 (**101,** Part III, p. 238).

(3) KARBOWIAK, A. E.: 'Microwave Propagation in Anisotropic Waveguides', *ibid.*, Monograph No. 147 R, August, 1955 (**103** C, p. 139).

(4) KARBOWIAK, A. E.: 'The Concept of Heterogeneous Surface Impedance and its Application to Cylindrical Cavity Resonators', *ibid.*, Monograph No. 246 R, June, 1957 (**105** C).

(5) KARBOWIAK, A. E.: 'Microwave Aspects of Waveguides for Long-Distance Transmission', presented at the Congrès Internationale Circuits et Antennes Hyperfréquences, Paris, October, 1957 (to be published).

621.317.335.3: 621.317.41: 621.318.12

The Institution of Electrical Engineer
Paper No. 2518 F
Mar. 195

MEASUREMENT OF FERRITE LOSS-FACTORS AT 10 Gc/s

By C. M. SRIVASTAVA, M.Sc., Ph.D., and J. ROBERTS, B.Sc.(Eng.), Associate Member.

(*The paper was first received 24th September, and in revised form 19th November,* 1957.)

SUMMARY

The analysis is presented of the losses arising in a rectangular cavity containing ferrite slabs which extend the full length of the cavity. The loss factors associated with the dielectric constant and the scalar and tensor permeabilities are deduced from Q-factor measurements on the cavity. The methods are particularly suited to low-loss ferrites.

LIST OF SYMBOLS

μ_s = Complex scalar permeability of ferrite, henrys/m.
μ_r' and μ_r'' = Real and imaginary parts respectively of the complex relative scalar permeability.
μ_0 = Permeability of free space, henrys/m.
$\mu = \mu' - j\mu''$ = Complex diagonal term of relative permeability tensor.
$K = K' - jK''$ = Complex off-diagonal term of the relative permeability tensor.
ϵ = Complex dielectric constant of the ferrite, F/m.
ϵ_r' and ϵ_r'' = Real and imaginary parts respectively of the complex relative dielectric constant.
ϵ_0 = Dielectric constant of free space, F/m.
D = Electric flux density, coulombs/m^2.
B = Magnetic flux density, Wb/m^2.
β = Phase constant, rad/m.
k_a = Transverse phase constant in air, rad/m.
k_m = Transverse phase constant in the ferrite, rad/m.
$g(x) = x - \sin x$.
$h(x) = x + \sin x$.
F = Frequency, Mc/s.

(1) INTRODUCTION

The loss factors associated with the microwave use of ferrites have been determined previously[1–5] from measurements whose interpretation involves first-order perturbation theory. This restricts the size of the ferrite sample which may be used and seriously limits the accuracy of loss determinations in the case of low-loss ferrites such as Mg–Mn ferrite.

The methods described here are based on the exact solution of the field equations which can be found for one or more slabs of ferrite in a rectangular waveguide and there is no severe limitation of slab thickness. The cavity techniques adopted for the measurement of the real parts of the scalar and tensor properties have been described previously.[6] The analysis is extended here to the interpretation of the measured Q-factor of the cavity in terms of the imaginary or loss component of the scalar and tensor properties.

The determination of the real and imaginary parts of both the dielectric constant and the scalar permeability involves the measurement of the shift in resonant frequency, and of the Q-factor of the cavity, when the ferrite slab is placed first along the side wall and secondly along the centre-line of the cavity. These two locations are chosen since the measurements are mainly dependent on the permeability in the first arrangement and the dielectric constant in the second arrangement.

The tensor properties of the ferrite are determined from the shift in resonance frequency and the Q-factor of the cavity when the latter is symmetrically loaded with two identical slabs of ferrite, one along each side wall, the ferrite slabs being magnetized in the same direction. A second set of measurements with slabs of different thickness complete the data necessary to specify the tensor properties.

(2) COMPLEX SCALAR PERMEABILITY AND DIELECTRIC CONSTANT

(2.1) Theory

The losses associated with the behaviour of materials in low-level magnetic and electric fields are generally incorporated in the permeability and dielectric constant by assigning complex values to these quantities. Thus, in an isotropic medium, the electric flux density D and magnetic flux density B are related to the electric field strength E and magnetic field strength H, respectively, by

$$\left.\begin{aligned} D &= \epsilon E = \epsilon_0(\epsilon_r' - j\epsilon_r'')E \\ \text{and} \quad B &= \mu_s H = \mu_0(\mu_r' - j\mu_r'')H \end{aligned}\right\} \quad . \quad . \quad . \quad (1)$$

Since $\epsilon_r'' \ll \epsilon_r'$ and $\mu_r'' \ll \mu_r'$ in the case of the low-loss ferrites being considered here, the field distribution in the cavity and the stored energy in the electric and magnetic fields will be deduced on the assumption that ϵ and μ_s are purely real.

(2.1.1) Determination of μ_r' and ϵ_r'.

In the case of a rectangular waveguide containing an unmagnetized ferrite slab, shown in Fig. 1, the solution of Maxwell's

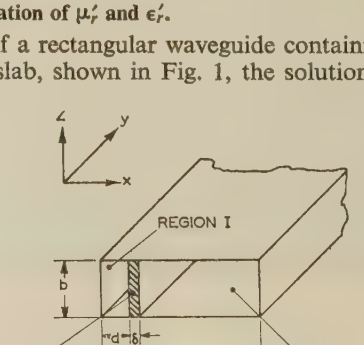


Fig. 1.—Rectangular waveguide containing an unmagnetized ferrite slab.

equations for a TE-mode propagating in the y-direction with a phase constant β yields the following distribution of electric field strength in the three regions:

$$\left.\begin{aligned} &\text{Region I.} \quad 0 \leqslant x \leqslant d\text{: } E_z = A \sin k_a x \exp(-j\beta y) \\ &\text{Region II.} \quad d \leqslant x \leqslant d + \delta\text{: } E_z = (C \sin k_m x \\ &\qquad\qquad + D \cos k_m x) \exp(-j\beta y) \\ &\text{Region III.} \quad d + \delta \leqslant x \leqslant L\text{:} \\ &\qquad\qquad E_z = B \sin k_a(L - x) \exp(-j\beta y) \end{aligned}\right\} \quad (2)$$

Written contributions on papers published without being read at meetings are invited for consideration with a view to publication.

Dr. Srivastava was formerly and Mr. Roberts is in the Department of Electrical Engineering, Imperial College of Science and Technology, University of London.
Dr. Srivastava is now at the College of Science, Benares Hindu University, India.

The time factor $\exp(j\omega t)$ is assumed throughout and the transverse phase constants, k_a and k_m, in the air and ferrite regions, respectively, are given by

$$\left.\begin{aligned} k_a^2 &= \omega^2\epsilon_0\mu_0 - \beta^2 \\ \text{and}\quad k_m^2 &= \omega^2\epsilon_r'\epsilon_0\mu_r'\mu_0 - \beta^2 \end{aligned}\right\} \quad . \quad . \quad . \quad (3)$$

The longitudinal components of magnetic field strength derived from eqns. (2) are as follows:

$$\left.\begin{aligned} &\text{Region I.} \quad H_y = -\left(\frac{jk_aA}{\omega\mu_0}\right)\cos k_a x \exp(-j\beta y) \\ &\text{Region II.} \quad H_y = -\left(\frac{jk_m}{\omega\mu_0\mu_r'}\right)(C\cos k_m x - D\sin k_m x)\exp(-j\beta y) \\ &\text{Region III.} \quad H_y = \left(\frac{j\beta k_a}{\omega\mu_0}\right)\cos k_a(L-x)\exp(-j\beta y) \end{aligned}\right\} \quad (4)$$

Since the tangential components of both the electric and magnetic field strengths must be continuous at the air–ferrite interface, the following boundary conditions must be satisfied:

$$\left.\begin{aligned} A\sin k_a d &= C\sin k_m d + D\cos k_m d \\ \frac{Ak_a}{\mu_0}\cos k_a d &= \frac{k_m}{\mu_0\mu_r'}(C\cos k_m d - D\sin k_m d) \\ B\sin k_a(L-d-\delta) &= C\sin k_m(d+\delta) + D\cos k_m(d+\delta) \\ \frac{k_a}{\mu_0}\cos k_a(L-d-\delta) &= \frac{k_m}{\mu_0\mu_r'}[D\sin k_m(d+\delta) - C\cos k_m(d+\delta)] \end{aligned}\right\} \quad (5)$$

Eliminating the arbitrary amplitude constants A, B, C and D yields

$$(\mu_r'^2k_a^2 - k_m^2)\cos k_a(L-\delta-2d) + (\mu_r'^2k_a^2 + k_m^2)\cos k_a(L-\delta) + 2\mu_r'k_ak_m\cot k_m\delta\sin k_a(L-\delta) = 0 \quad (6)$$

There are two cases of interest here. First, with the ferrite slab in contact with the side wall, i.e. $d = 0$; secondly with the slab along the centre-line of the guide, i.e. $d = 1/2(L-\delta)$.

Case I. $d = 0$.

For this case, the transcendental equation reduces to

$$\frac{\tan k_m\delta}{k_m\delta} = -\frac{1}{\mu_r'}\frac{\tan k_a(L-\delta)}{k_a\delta} \quad . \quad . \quad . \quad (7)$$

Case II. $d = 1/2(L-\delta)$.

In this case, the transcendental equation reduces to

$$\frac{k_m^2 d}{2}\frac{\tan k_m\delta/2}{k_m\delta/2} = \mu_r'k_a\cot 1/2k_a(L-\delta) \quad . \quad . \quad (8)$$

Two experimental determinations of β will suffice to find μ_r' and ϵ_r'; either (*a*) using two different thicknesses, δ, of ferrite slab in contact with the side wall and solving eqn. (7), or (*b*) using the same slab, first in contact with the side wall, and then along the centre-line of the waveguide and solving eqns. (7) and (8) for the appropriate arrangements.

2) Determination of μ_r'' and ϵ_r''.

The losses arising in a cavity containing a ferrite slab are due partly to the magnetic and dielectric losses in the ferrite and partly to the wall losses in the cavity, which will depend on the field distribution and hence on the ferrite location. In order to determine the imaginary components μ_r'' and ϵ_r'', it is necessary to relate these to the experimentally observable Q-factor of the cavity for the two arrangements considered above.

(*a*) *Ferrite Slab in Contact with Side-Wall of Cavity.*

Eqns. (2) for the electric-field distribution in a waveguide containing a ferrite slab in contact with the side wall reduce to

$$\left.\begin{aligned} &\text{Region II.} \quad 0 \leqslant x \leqslant \delta\text{: } E_z = C\sin k_m x\exp(-j\beta y) \\ &\text{Region III.} \quad \delta \leqslant x \leqslant L\text{: } E_z = B\sin k_a(L-x)\exp(-j\beta y) \end{aligned}\right\} \quad (9)$$

If a cavity is formed by closing the waveguide with short-circuiting plates at the ends of the ferrite slab of length l, then the cavity will be resonant at a frequency such that $\beta = n\pi/l$, where n is an integer. Further, if one short-circuiting plate is in the plane $y = 0$, the electric-field distribution in the cavity will be given by

$$\left.\begin{aligned} &\text{Region II.} \quad E_z = 2C\sin k_m x\sin\beta y \\ &\text{Region III.} \quad E_z = 2B\sin k_a(L-x)\sin\beta y \end{aligned}\right\} \quad . \quad (10)$$

The corresponding magnetic-field components are

$$\left.\begin{aligned} &\text{Region II.} \quad H_x = \frac{2jC\beta}{\omega\mu_0\mu_r'}\sin k_m x\cos\beta y \\ &\qquad\qquad H_y = -\frac{2jCk_m}{\omega\mu_0\mu_r'}\cos k_m x\sin\beta y \\ &\text{Region III.} \quad H_x = \frac{2jB\beta}{\omega\mu_0}\sin k_a(L-x)\cos\beta y \\ &\qquad\qquad H_y = \frac{2jBk_a}{\omega\mu_0}\cos k_a(L-x)\sin\beta y \end{aligned}\right\} \quad . \quad (11)$$

The stored energy, W, in the cavity may be expressed as the peak stored electric energy

$$W_E = \tfrac{1}{2}\int_V \epsilon_0\epsilon_r'\hat{E}^2 dV \quad . \quad . \quad . \quad . \quad (12)$$

or as the peak stored magnetic energy,

$$W_M = \tfrac{1}{2}\int_V \mu_0\mu_r'\hat{H}^2 dV \quad . \quad . \quad . \quad . \quad (13)$$

where the integration is carried out throughout the volume, V, enclosed by the cavity.

From eqns. (10) and (12), it follows that

$$W = \frac{\epsilon_0\epsilon_r' blC^2}{4k_m}g(2k_m\delta) + \frac{\epsilon_0 blB^2}{4k_a}g[2k_a(L-\delta)] \quad . \quad (14)$$

where l is the length and b the width of the cavity; and $g(x) = [x - \sin x]$.

The power dissipation, P_s, in the specimen due to magnetic and dielectric losses is given by

$$P_s = \frac{\omega}{2}\int_{V'}(\epsilon_r''\epsilon_0\hat{E}^2 + \mu_r''\mu_0\hat{H}^2)dV \quad . \quad . \quad . \quad (15)$$

where the integration extends over the volume, V', occupied by the ferrite slab. Making appropriate substitutions in eqn. (15), it may be shown that

$$P_s = \frac{\omega\epsilon_r''\epsilon_0 blC^2}{4k_m}g(2k_m\delta) + \frac{\mu_r'' blC^2}{4k_m\omega(\mu_r')^2\mu_0}[k_m^2h(2k_m\delta) + \beta^2 g(2k_m\delta)] \quad (16)$$

where $h(x) = (x + \sin x)$

The power dissipation, P_w, in the walls of the cavity is given by

$$P_w = \tfrac{1}{2}R_s \int_s \hat{H}_t^2 dS \quad . \quad . \quad . \quad . \quad (17)$$

where R_s = Resistance per unit area of conductor surface.
H_t = Component of magnetic field intensity tangential to the wall of the cavity.

The integration is carried out over the entire area S of the inner surface of the cavity.

Substituting from eqn. (11) into (17), it is found that

$$P_w = \frac{R_s}{\omega^2\mu_0^2}\left[\frac{C^2}{(\mu_r')^2}\left[lbk_m^2 + \frac{(2b+l)\beta^2}{2k_m}g(2k_m\delta) + \frac{k_m l}{2}h(2k_m\delta)\right] + B^2\left\{lbk_a^2 + \frac{(2b+l)\beta^2}{2k_a}g[2k_a(L-\delta)] + \frac{k_a l}{2}h[2k_a(L-\delta)]\right\}\right] \quad . \quad . \quad . \quad . \quad (18)$$

The Q-factor of the cavity is given by

$$Q = \frac{\omega W}{P_s + P_w} \quad . \quad . \quad . \quad . \quad . \quad (19)$$

where W, P_s and P_w are given by eqns. (14), (16) and (18) respectively.

The relation between the amplitude constants B and C can be obtained by equating E_z for the two regions at the air–ferrite interface, giving

$$B = C\frac{\sin k_m\delta}{\sin k_a(L-\delta)} \quad . \quad . \quad . \quad . \quad (20)$$

Hence, finally, eqn. (21) is obtained

$$Q = \frac{\dfrac{\epsilon_r'}{k_m}g(2k_m\delta) + \dfrac{\sin^2 k_m\delta}{k_a \sin^2 k_a(L-\delta)}g[2k_a(L-\delta)]}{\dfrac{\epsilon_r''}{k_m}g(2k_m\delta) + \dfrac{2\,280\mu_r''}{k_m(\mu_r')^2F^2}[k_m^2 h(2k_m\delta) + \beta^2 g(2k_m\delta)] + \dfrac{1\,150R_s}{blF^3}\left[\dfrac{1}{(\mu_r')^2}\left[k_m^2 bl + \dfrac{(2b+l)\beta^2}{2k_m}g(2k_m\delta) + \dfrac{k_m l}{2}h(2k_m\delta)\right] + \dfrac{\sin^2 k_m\delta}{\sin^2 k_a(L-\delta)}\left\{k_a^2 bl + \dfrac{(2b+l)\beta^2}{2k_a}g[2k_a(L-\delta)] + \dfrac{k_a l}{2}h[2k_a(L-\delta)]\right\}\right]} \quad . \quad . \quad . \quad . \quad (21)$$

(*b*) *Ferrite Slab along Centre-Line of Cavity.*

Proceeding along similar lines to those in the previous Section, eqn. (22) is obtained for this case

$$Q = \frac{\dfrac{\epsilon_r'}{k_m}h(k_m\delta) + \dfrac{\cos^2(k_m\delta/2)}{k_a \sin^2[k_a(L-\delta)/2]}g[k_a(L-\delta)]}{\dfrac{\epsilon_r''}{k_m}h(k_m\delta) + \dfrac{2\,280\mu_r''}{k_m(\mu_r')^2F^2}[k_m^2 g(k_m\delta) + \beta^2 h(k_m\delta)] + \dfrac{1\,150R_s}{blF^3}\left[\dfrac{1}{(\mu_r')^2}\left[\dfrac{(2b+l)\beta^2}{2k_m}h(k_m\delta) + \dfrac{k_m l}{2}g(k_m\delta)\right] + \dfrac{\cos^2(k_m\delta/2)}{\sin^2[k_a(L-\delta)/2]}\left\{k_a^2 bl + \dfrac{(2b+l)\beta^2}{2k_a}g[k_a(L-\delta)] + \dfrac{k_a l}{2}h[k_a(L-\delta)]\right\}\right]} \quad . \quad . \quad . \quad . \quad (22)$$

(2.2) Experimental Procedure

The cavity used in these experiments is formed from waveguide No. 16 by closing one end with a short-circuiting plate which has a circular iris for coupling to a microwave test bench, and by closing the other end with a short-circuiting plunger to provide flexibility. In every experiment the ferrite slab extends the full length of the cavity to avoid end-effects. The test bench includes a directional coupler to facilitate location of the cavity resonant frequency by direct observation of the signal reflected from the cavity, and also a standing-wave detector.

The procedure is as follows: with the ferrite slab in contact with the side wall of the cavity the variation with frequency of the standing-wave ratio in the input guide to the cavity is measured. From these measurements the resonant frequency and the unloaded Q-factor of the cavity can be deduced by well-known methods. The ferrite slab is then located along the centre-line of the cavity and the measurements and deductions repeated; in this second part it is usually more convenient to observe the TE_{012}-mode, rather than the TE_{011}-mode examined in the first part, since β is usually appreciably greater. Finally the resonant frequency and Q-factor of the empty cavity are determined and the value of R_s, the surface resistance of the wall of the cavity at this frequency, deduced; the value of R_s at the other resonant frequencies can than be deduced since it is proportional to the square root of the frequency.

(2.3) Experimental Observations and Results

The measurements reported here were carried out using magnesium-manganese ferrite, S.E.R.L. 1009; the recommendations made by Blackman[7] concerning die pressure and firing temperature were followed in forming the test slabs. The observations with a slab along the side wall were:

(*a*) Thickness of the ferrite slab, δ, = 0·110 5 cm.
(*b*) Length of the slab, l, = 2·311 cm.
(*c*) Resonant frequency of the empty cavity = 9·169 4 Gc/s.
(*d*) Q-factor of the empty cavity = 3 400.
(*e*) Resonant frequency of the ferrite-loaded cavity = 9·160 5 Gc/s.
(*f*) Q-factor of the ferrite-loaded cavity = 1 500.

The same slab was then located along the centre-line and the following observations were made.

(*g*) Resonant frequency of the ferrite-loaded cavity = 9·648 Gc/s.
(*h*) Q-factor of the ferrite-loaded cavity = 790.

The values of ϵ_r', ϵ_r'', μ_r' and μ_r'' calculated from these measurements are

$$\epsilon = \epsilon_0(\epsilon_r' - j\epsilon_r'') = \epsilon_0(10{\cdot}0 - j1{\cdot}1 \times 10^{-3})$$
$$\mu = \mu_0(\mu_r' - j\mu_r'') = \mu_0(0{\cdot}91 - j5{\cdot}4 \times 10^{-3})$$

The thickness could be measured with an accuracy of 0·5%, d the error due to it in the final values of ϵ and μ is n 2% for the real parts, ϵ_r' and μ_r', and less than 5% the imaginary parts, ϵ_r'' and μ_r''. The Q-factor could be asured with an accuracy of 2%, and the error due to it is out 10% for ϵ_r'' and 4% for μ_r''. The frequency difference uld be measured with an accuracy of 0·1 Mc/s, and this roduces negligible error in the final values of ϵ and μ.

(3) COMPLEX TENSOR PERMEABILITY

(3.1) Theory

When a ferrite is magnetized by the application of a steady gnetic field in the z-direction, the magnetic flux densities in x- and y-directions are found each to depend on both the and y-components of an applied high-frequency alternating gnetic field. This may be represented by assigning to the rite a tensor permeability of the form

$$\mu_0 \begin{vmatrix} \mu & -jK & 0 \\ jK & \mu & 0 \\ 0 & 0 & 1 \end{vmatrix} \quad . \quad . \quad . \quad . \quad . \quad . \quad (23)$$

As in the case of the scalar permeability, losses may be repre- ted by allowing the tensor components μ and K to take on mplex values. Thus

$$\left.\begin{aligned} \mu &= \mu' - j\mu'' \\ \text{d} \qquad K &= K' - jK'' \end{aligned}\right\} \quad . \quad . \quad . \quad . \quad . \quad . \quad (24)$$

Of the two methods described elsewhere[6] for determining and K', the twin-slab reciprocal cavity has been chosen for ther study of the loss components μ'' and K'', since in this e the calculations are not complicated by duality of the phase nstant.

The transverse electric-field distribution, for propagation in positive y-direction in a rectangular waveguide carrying two tical ferrite slabs in contact with its side walls, the ferrite bs being magnetized in the positive z-direction, may be itten[8] (see Fig. 2).

$$\left.\begin{aligned} &\text{Region I.} \quad 0 \leqslant x \leqslant \delta \qquad E_z = A \sin k_m x \exp(-j\beta y) \\ &\text{Region II.} \quad \delta \leqslant x \leqslant L-\delta \quad E_z = (C \cos k_a x \\ &\qquad\qquad + D \sin k_a x) \exp(-j\beta y) \\ &\text{Region III.} \quad L-\delta \leqslant x \leqslant L \quad E_z = B \sin k_m \\ &\qquad\qquad (L-x) \exp(-j\beta y) \end{aligned}\right\} (25)$$

The time factor $\exp(j\omega t)$ is assumed throughout and the nsverse phase constants k_a and k_m are given by

$$\left.\begin{aligned} k_a^2 &= \omega^2\mu_0\epsilon_0 - \beta^2 \\ \text{d} \qquad k_m^2 &= \omega^2\mu_0\epsilon_0\epsilon_r' \frac{[(\mu')^2 - (K')^2]}{\mu'} - \beta^2 \end{aligned}\right\} \quad (26)$$

Owing to the tensor permeability of the ferrite, the amplitude nstants A and B in these expressions are not equal. For opagation in the negative y-direction, the phase constant will the same as that given above for the positive y-direction of opagation since the ferrite is symmetrically disposed and the ady magnetic field is unidirectional. The field distribution

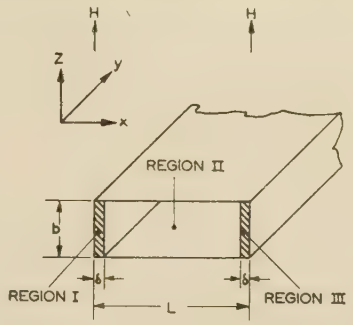


Fig. 2.—Ferrite-loaded waveguide.

for the negative y-direction of propagation will also be similar to that for the positive y-direction of propagation when each is viewed with respect to its direction of propagation. The two field distributions will, therefore, be mirror images of each other in the plane $x = L/2$, but since $A \neq B$, the amplitudes will not be equal at any arbitrary point in the waveguide. If a ferrite-loaded waveguide such as shown in Fig. 2 is terminated in a short-circuit, this non-correspondence of field patterns will give rise to other modes at the short-circuit. However, these modes will be evanescent and, therefore, they will be restricted to the immediate vicinity of the short-circuit.

Since the amplitudes A and B in regions I and III are merely interchanged for the two directions of propagation, the problem of energy storage and power loss in a cavity symmetrically loaded with ferrite can be treated as though the transverse electric-field distribution of eqns. (25) were reflected in anti-phase at the short-circuited ends of the cavity, provided that the energy associated with the evanescent modes is small compared with the total energy stored in the cavity. This latter condition requires that A be not very different from B, and it is dependent on the thickness of the ferrite slabs; the agreement between results obtained for various values of slab thickness up to 2·1 mm suggests that this disregard of the evanescent modes is justified for thicknesses not exceeding this value.

With these simplifications, the electric field distribution in the cavity at resonance is given by

$$\left.\begin{aligned} &\text{Region I.} \quad E_z = 2A \sin k_m x \sin \beta y \\ &\text{Region II.} \quad E_z = 2(C \cos k_a x + D \sin k_a x) \sin \beta y \\ &\text{Region III.} \quad E_z = 2B \sin k_m (L-x) \sin \beta y \end{aligned}\right\} \quad (27)$$

The magnetic field distribution is similarly compounded from the modified patterns of the forward and reverse waves.

Making appropriate substitutions in eqn. (12), the peak stored electric energy, W_E, may be shown to be

$$W_E = \frac{\epsilon_0 bl}{4k_a}\{2k_a(C^2+D^2)(L-2\delta) + (C^2-D^2)[\sin 2k_a(L-\delta) - \sin 2k_a\delta] + 2CD[\cos 2k_a\delta - \cos 2k_a(L-\delta)]\}$$
$$+ \frac{\epsilon_0\epsilon_r' bl}{4k_m}(A^2+B^2)g(2k_m\delta) \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (28)$$

Since the x- and y-components of the magnetic field intensity H are in phase with the corresponding components of magnetic flux density, B, in each region, the peak stored magnetic energy W_M is given by

$$W_M = \tfrac{1}{2}\int_V (\hat{H}_x\hat{B}_x + \hat{H}_y\hat{B}_y)dV$$
$$= \frac{bl}{4k_a\mu_0\omega^2}\{2k_a(C^2+D^2)(L-2\delta)(k_a^2+\beta^2) - 2CD(k_a^2-\beta^2)[\cos 2k_a\delta - \cos 2k_a(L-\delta)] + (D^2-C^2)(k_a^2-\beta^2)[\sin 2k_a(L-\delta) - \sin 2k_a\delta]\}$$
$$+ \frac{bl}{4k_m\mu_0\omega^2[(\mu')^2-(K')^2]}[2k_m\delta\mu'(A^2+B^2)(k_m^2+\beta^2) + \mu'(A^2+B^2)(k_m^2-\beta^2)\sin 2k_m\delta + 2\beta k_m K'(A^2-B^2)(1-\cos 2k_m\delta)] \quad (29)$$

By applying the boundary conditions at the air-ferrite interface and using eqns. (28) and (29), it may be shown that $W_E = W_M$, and either may be used as the stored energy, W, in the cavity.

The dielectric and magnetic losses, P_s, in the ferrite are given by

$$P_s = \frac{\omega\epsilon_0\epsilon_r'' bl}{4k_m}(A^2 + B^2)g(2k_m\delta) + \frac{bl}{4k_m\mu_0\omega[(\mu')^2 - (K')^2]}\left[2k_m\delta\mu''(A^2 + B^2)(k_m^2 + \beta^2) + \mu''(A^2 + B^2)(k_m^2 - \beta^2)\sin 2k_m\delta + 2\beta k_m K''(A^2 - B^2)(1 - \cos 2k_m\delta)\right] \quad . \quad . \quad . \quad . \quad (30)$$

The wall losses, P_W, in the ferrite loaded cavity are

$$P_W = \frac{R_s}{2k_a\omega^2\mu_0^2}\left\{(C^2 + D^2)2k_a(L - 2\delta)[(2b + l)\beta^2 + lk_a^2] + (C^2 - D^2)[\sin 2k_a(L - \delta) - \sin 2k_a\delta][(2b + l)\beta^2 - lk_a^2] + 2CD[\cos 2k_a\delta - \cos 2k_a(L - \delta)][(2b + l)\beta^2 - lk_a^2]\right\} + \frac{R_s}{2k_m\omega^2\mu_0^2[(\mu')^2 - (K')^2]^2}\left[(A^2 + B^2)\left(2k_m^3 bl(\mu')^2 + k_m^2\left\{2b(K')^2 + l[(\mu')^2 + (K')^2]\right\}h(2k_m\delta) + \beta^2\left\{2b(\mu')^2 + l[(\mu')^2 + (K')^2]\right\}g(2k_m\delta)\right) + 4(A^2 - B^2)(b + l)\mu' K'\beta k_m(1 - \cos 2k_m\delta)\right] \quad (31)$$

Collecting terms,

$$Q = \frac{\omega W}{P_S + P_W} = \frac{\frac{bl}{k_a}\left\{2k_a(C^2 + D^2)(L - 2\delta) + (C^2 - D^2)[\sin 2k_a(L - \delta) - \sin 2k_a\delta] + 2CD[\cos 2k_a\delta - \cos 2k_a(L - \delta)]\right\} + \frac{\epsilon_r' bl}{k_m}(A^2 + B^2)g(2k_m\delta)}{\frac{\epsilon_r'' bl}{k_m}(A^2 + B^2)g(2k_m\delta) + \frac{2\,280bl}{k_m F^2[(\mu')^2 - (K')^2]}\left\{\mu''(A^2 + B^2)[(k_m^2 + \beta^2)2k_m\delta + (k_m^2 - \beta^2)\sin 2k_m\delta] + K''(A^2 - B^2)\,2\beta k_m(1 - \cos 2k_m\delta)\right\} + \frac{575\,R_s}{k_a F^3}\left\{(C^2 + D^2)2k_a(L - 2\delta)[(2b + l)\beta^2 + lk_a^2] + (C^2 - D^2)[\sin 2k_a(L - \delta) - \sin 2k_a\delta][(2b + l)\beta^2 - lk_a^2] + 2CD[\cos 2k_a\delta - \cos 2k_a(L - \delta)][(2b + l)\beta^2 - lk_a{}^2]\right\} + \frac{575\,R_s}{k_m F^3[(\mu')^2 - (K')^2]^2}\left[(A^2 + B^2)\left(2k_m^3 bl(\mu')^2 + k_m^2 h(2k_m\delta) \times \left\{2b(K')^2 + l[(\mu')^2 + (K')^2]\right\} + \beta^2 g(2k_m\delta) \times \left\{2b(\mu')^2 + l[(\mu')^2 + (K')^2]\right\}\right) + 4(A^2 - B^2)(b + l)\mu' K'\beta k_m(1 - \cos 2k_m\delta)\right]} \quad (32)$$

In this expression the amplitude constants A, B, C and D are related by imposing the boundary conditions at the air–ferrite interfaces, i.e.

$$A\sin k_m\delta = C\cos k_a\delta + D\sin k_a\delta$$
$$B\sin k_m\delta = C\cos k_a(L - \delta) + D\sin k_a(L - \delta)$$

and $A(\mu' k_m\cos k_m\delta + \beta K'\sin k_m\delta) = k_a[(\mu')^2 - (K')^2](D\cos k_a\delta - C\sin k_a\delta) \quad (33)$

These three conditions suffice to eliminate the four amplitude constants from eqn. (32).

(3.2) Experimental Procedure and Results

In this method, the frequency of resonance and unloaded Q-factor of a rectangular TE_{011} cavity containing two transversely magnetized ferrite slabs resting against its side walls and completely filling the cavity along its length and narrow dimension are obtained by the same method as described in Section 2.2. Since eqn. (32) contains two unknowns, the experiment is repeated for a different thickness of the slabs. From these two measurements μ'' and K'' can be deduced using the values of R_s and ϵ'' obtained in Section 2.

In a typical measurement with two magnesium-manganese ferrite slabs, the following observations were made:

(*a*) Thickness of the ferrite slabs = 0·106 5 cm.
(*b*) Length of the ferrite slabs = 2·161 cm.
(*c*) Resonant frequency of the empty cavity = 9·466 Gc/s.
(*d*) Q-factor of the empty cavity = 3 400.
(*e*) Applied field = 250 oersteds.
(*f*) Resonant frequency of the ferrite-loaded cavity = 9·508 8 Gc/s.
(*g*) Q-factor of the ferrite-loaded cavity = 1 740.

The thickness of the ferrite was then changed and the following observations made:

(*h*) Thickness of the ferrite slab = 0·085 cm.
(*i*) Resonant frequency of the ferrite-loaded cavity = 9·529 9 Gc/s.
(*j*) Q-factor of the ferrite-loaded cavity = 2 270.

The values of μ', μ'', K' and K'' calculated from these measurements using the value of 10·0 for ϵ' and $1{\cdot}1 \times 10^{-3}$ for ϵ'' are

$$\mu = \mu' - j\mu'' = 0{\cdot}98 - j2{\cdot}0 \times 10^{-3}$$
$$K = K' - jK'' = 0{\cdot}45 - j17 \times 10^{-3}$$

The thickness of the slab could be measured with an accuracy of 0·5%, and the error in the final values of μ and K due to this is less than 2% for both the real and imaginary parts. The Q-factor of the cavity could be measured with an accuracy of 2%, and this introduces an error of less than 10% in the values of μ'' and K''. The error due to the inaccuracy in frequency determination is negligible for all the four quantities.

(4) CONCLUSION

The methods used for the measurement of the imaginary parts of μ and K give results of much better accuracy than those obtained by perturbation techniques for low-loss ferrites. Macbean[9] reports an error of 100% for his measurements of the imaginary parts of μ and K for magnesium-manganese ferrite, while for the same measurements using this technique the corresponding error is only about 15%.

Although the regions near resonance cannot be investigated with this technique, it is of value in obtaining accurate ferrite parameters for the development of the many microwave devices which work in regions away from resonance.

With magnesium-manganese ferrite, measurements have been made up to an applied steady magnetic field of approximately 2 000 oersteds.

(5) ACKNOWLEDGMENT

This paper is published by permission of the Admiralty.

(6) REFERENCES

ARTMAN, J. O., and TANNENWALD, P. E.: 'Measurement of Permeability Tensor in Ferrites', *Physical Review*, 1953, **91**, p. 1014.

'Measurement of Susceptibility Tensor in Ferrites', *Journal of Applied Physics*, 1955, **26**, p. 1124.

SPENCER, E. G., LEGRAW, R. C., and REGGIA, F.: 'Circularly Polarized Cavities for Measurement of Tensor Permeabilities', *ibid.*, 1955, **26**, p. 354.

'Measurement of Microwave Dielectric Constants and Tensor Permeabilities of Ferrite Spheres', *Institute of Radio Engineers*, Convention Record, 1955, Part 8, p. 113.

SPENCER, E. G., AULT, L. A., and LEGRAW, R. C.: 'Intrinsic Tensor Permeabilities of Ferrite Rods, Spheres and Discs', *Proceedings of the Institute of Radio Engineers*, 1956, **44**, p. 1311.

MACBEAN, I. G.: 'The Measurement of Complex Permittivity and Complex Tensor Permeability of Ferrite Materials at Microwave Frequencies', *Proceedings I.E.E.*, Paper No. 2205, October, 1956 (**104** B, Supplement No. 6, p. 296).

(5) VON AULOCK, W., and ROWEN, J. H.: 'Measurement of Dielectric and Magnetic Properties of Ferromagnetic Materials at Microwave Frequencies', *Bell System Technical Journal*, 1957, **36**, p. 427.

(6) ROBERTS, J., and SRIVASTAVA, C. M.: 'Measurement of Ferrite Properties in a Rectangular Cavity at 10 000 Mc/s', *Proceedings I.E.E.*, Paper No. 2219 R, October, 1956 (**104** B, Supplement No. 6, p. 338).

(7) BLACKMAN, L. C. F.: 'Low Loss Magnesium Manganese Ferrites', *Journal of Electronics*, 1957, **2**, p. 451.

(8) LAX, B., and BUTTON, K. J.: 'Theory of New Ferrite Modes in Rectangular Wave Guide', *Journal of Applied Physics*, 1955, **26**, p. 1184.

(9) MACBEAN, I. G.: Ph.D. Thesis presented to the University of London, 1956.

621.372.852.1

The Institution of Electrical Engineers
Paper No. 2547 R
Mar. 1958

WIDE-BAND WAVEGUIDE FILTERS WITH SHORT LINEAR TAPERS

By G. CRAVEN.

(*The paper was first received 27th September, and in revised form 30th November*, 1957.)

SUMMARY

Broad-band waveguide filters, which obtain a band-pass characteristic by combining low-pass and high-pass networks, are discussed. Filters of this type do not, in general, match the impedance of standard guide. A method of design, in which simple linear tapers are employed to effect matching, is described. The design procedure is illustrated in a typical example and experimental results are given.

(1) INTRODUCTION

When waveguide band-pass filters are required for bandwidths exceeding about 2% of mid-band frequency, it is common practice to employ the type described by Cohn,[1] in which high-pass and low-pass filters are combined to provide the band-pass characteristic. The high-pass filter consists simply of a section of plain guide, the width being chosen to give the desired cut-off frequency, with the overall length as the factor controlling band-stop attenuation. The low-pass filter is made up of alternate sections of common width, but different narrow dimension, the cut-off frequency depending on the length and characteristic impedance of these sections. The desired band-stop attenuation is obtained by employing a sufficient number of sections. It is usual to construct the low-pass filter in the guide forming the high-pass network, since this leads to a more compact design. Cohn gives full design details and describes suitable waveguide-to-coaxial transducers. However, it is often necessary to employ these filters as part of a waveguide system and it is then desirable that they should be matched to standard guide. This can prove difficult in practice because the impedance of these filters can differ greatly from that of standard guide, necessitating extremely long tapers. This problem can be solved by modifying the design procedure somewhat and employing a particular kind of taper.

(2) FILTER DESIGN

The first step in filter design, with the normal procedure, is to calculate the guide width required for the desired high-pass cut-off frequency. With the low-pass and high-pass units constructed integrally this then determines the guide wavelength in the low-pass structure. The electrical lengths (at cut-off) of the sections forming the latter are then chosen so that the spurious pass-bands fall outside the band-stop region.[2] Using the guide wavelength at low-pass cut-off, the physical lengths can then be calculated. The choice of the electrical lengths of the sections also determines the ratio of the characteristic impedances, i.e. the ratio of the narrow dimensions of the two sections, if the cut-off conditions are to be satisfied. When this is found the actual dimensions required to satisfy this ratio may be decided, and, provided that this ratio is maintained, the only effect (within limits) of different dimensions is to change the image impedance of the filter. Thus, where the latter is not important a wide choice of dimensions is possible. However, where the image impedance is important, as when a taper must be designed to match the filter to standard guide, a modified procedure based on this parameter is more suitable. It then becomes possible to use a simple taper described by Lewin.[3]

In Lewin's paper a linear taper in both H- and E-planes is shown to be self-compensating if a particular relationship between the parameters holds. This occurs when

$$\frac{d-b}{D-a} = \frac{b}{a(1-\lambda_0^2/4a^2)} \quad \ldots\ldots \quad (1)$$

where a = Smaller wide dimension.
D = Larger wide dimension.
b = Smaller narrow dimension.
d = Larger narrow dimension.
λ_0 = Free-space wavelength.

Applying this to the present problem, a is the wide dimension chosen for the high-pass filter cut-off, D and d are, respectively, the wide and narrow dimensions of the standard guide, and, for self-compensation [from eqn. (1)],

$$b = \frac{d}{\dfrac{D-a}{a(1-\lambda_0^2/4a^2)} + 1} \quad \ldots\ldots \quad (1a)$$

The impedance at the contracted end of the taper can be found from b and a, and it is this impedance that the filter should be designed to match. Equations for the image impedance of a filter can be derived from the equivalent transmission-line sections. Denoting the characteristic impedances Z_{01} and Z_{02}

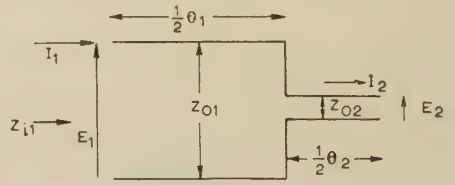


Fig. 1.—Schematic of filter half-section.

(where $Z_{01} > Z_{02}$), with their respective electrical lengths θ_1 and θ_2, the A-matrices in terms of half-sections (shown in Fig. 1) are

$$\begin{bmatrix} E_1 \\ I_1 \end{bmatrix} = \begin{bmatrix} \cos\frac{1}{2}\theta_1 & jZ_{01}\sin\frac{1}{2}\theta_1 \\ \frac{j}{Z_{01}}\sin\frac{1}{2}\theta_1 & \cos\frac{1}{2}\theta_1 \end{bmatrix} \times \begin{bmatrix} \cos\frac{1}{2}\theta_2 & jZ_{02}\sin\frac{1}{2}\theta_2 \\ \frac{j}{Z_{02}}\sin\frac{1}{2}\theta_2 & \cos\frac{1}{2}\theta_2 \end{bmatrix} \times \begin{bmatrix} E_2 \\ I_2 \end{bmatrix} \quad \ldots\ldots \quad (2)$$

The image impedance, Z_{i1}, can be obtained from the open- and short-circuit impedances of the network. In the A-matrix representing the network (see Fig. 1) this is given by

$$Z_{i1} = \sqrt{(Z_{OC}Z_{SC})} = \sqrt{\left(\frac{A_{11}}{A_{21}}\frac{A_{12}}{A_{22}}\right)}$$

Multiplying out in eqn. (2), we have

$$Z_{i1} = Z_{01}\left[\left(\frac{Z_{02} - Z_{01}\tan\frac{1}{2}\theta_1\tan\frac{1}{2}\theta_2}{Z_{02}\tan\frac{1}{2}\theta_1 + Z_{01}\tan\frac{1}{2}\theta_2}\right)\left(\frac{Z_{02}\tan\frac{1}{2}\theta_2 + Z_{01}\tan\frac{1}{2}\theta_1}{Z_{01} - Z_{02}\tan\frac{1}{2}\theta_1\tan\frac{1}{2}\theta_2}\right)\right]^{1/2} \quad (3)$$

Written contributions on papers published without being read at meetings are invited for consideration with a view to publication.

Mr. Craven is with Standard Telecommunication Laboratories Limited.

ie expression for the transfer constant can be obtained in a inilar way:

$$\cosh\frac{\gamma}{2} = \sqrt{(A_{11}A_{22})}$$

that

$$\text{lsh}\,\gamma = 2\left(\cos\tfrac{1}{2}\theta_1\cos\tfrac{1}{2}\theta_2 - \frac{Z_{01}}{Z_{02}}\sin\tfrac{1}{2}\theta_1\sin\tfrac{1}{2}\theta_2\right)$$

$$\times\left(\cos\tfrac{1}{2}\theta_1\cos\tfrac{1}{2}\theta_2 - \frac{Z_{02}}{Z_{01}}\sin\tfrac{1}{2}\theta_1\sin\tfrac{1}{2}\theta_2\right) - 1 \quad . \quad . \quad (4)$$

ut-off occurs when $\cosh\gamma = -1$ or $(Z_{01} > Z_{02})$ when

$$Z_{01}/Z_{02} = \cot\tfrac{1}{2}\theta_{1C}\cot\tfrac{1}{2}\theta_{2C} \quad . \quad . \quad . \quad . \quad (5)$$

ere the subscript C denotes cut-off.

These equations are used in the following way. The high-pass er is designed first in the normal manner and θ_{1C} and θ_{2C} are osen on the basis of considerations previously mentioned. The io Z_{01}/Z_{02} is obtained from eqn. (5) and the physical lengths θ_{1C} and θ_{2C} are determined. The electrical length at mid- nd, or wherever matching is to be obtained, is then calculated d the value substituted in eqn. (3). This gives the image pedance, KZ_{01} where K is a pure number. The characteristic pedance, Z_{01}, must then be such that the image impedance is al to guide of dimensions a, b, i.e. the contracted end of the er. The characteristic impedance, Z_0, of air-filled guide of se dimensions is[4]

$$Z_0 = 60\pi^2 b\left\{a[1 - (\lambda_0/2a)^2]^{1/2}\right\}^{-1} \quad . \quad . \quad . \quad (6)$$

milarly the impedance of the filter of guide dimensions a', b' is en by

$$KZ_{01} = K60\pi^2 b'\left\{a'[1 - (\lambda_0/2a')^2]^{1/2}\right\}^{-1} \quad . \quad . \quad (7)$$

Equating (6) and (7),

$$b'/b = a'[1 - (\lambda_0/2a')^2]^{1/2}\left\{Ka[1 - (\lambda_0/2a)^2]^{1/2}\right\}^{-1} \quad . \quad (8)$$

here $a' = a$ this simplifies to

$$b'/b = 1/K \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (9)$$

ith b' determined, the remaining dimension, b'', of the θ_2-section ay be found from the ratio Z_{01}/Z_{02}. Before completing the er design, corrections for the junction-discontinuity suscep- nces between the sections are made. The method follows that the original paper.[1]

Thus far, the assumption has been that the terminating half- ction is Z_{01}, θ_1 (where $Z_{01} > Z_{02}$). In many cases a Z_{02}, θ_2 rmination may be preferable. The lumped-circuit analogue of e Z_{01}, θ_1 termination is a T-section prototype filter, which has ro resistance at cut-off, whereas the Z_{02}, θ_2 termination, being uivalent to a π-section, is of infinite resistance. Thus, if match- g reasonably close to cut-off is desired the latter termination is ely to be better because it leads to larger values of K. The ysical dimensions b' and b'', which otherwise may be unde- ably large, are then reduced.

(3) DESIGN EXAMPLE

A filter with a pass-band of 3 800–4 200 Mc/s is required to atch 2 in × ¾ in guide. The attenuations at 3 500 Mc/s and 00 Mc/s should exceed 30 and 20 dB, respectively. Taking t-off at approximately 5% above and below the transmission nd, we have 3 600 Mc/s (8·33 cm) and 4 400 Mc/s (6·82 cm) as e cut-off frequencies. For the high-pass network cut-off urs when $\lambda_0 = 2a$, i.e. $a = 4{\cdot}17$ cm. Substituting for a, $D(=5{\cdot}08\,\text{cm})$ and $d(=1{\cdot}69\,\text{cm})$ in eqn. (1a), $b = 0{\cdot}79$ cm. The taper must be a linear one, contracting from 5·08 cm × 1·69 cm guide to 4·17 cm × 0·79 cm guide. The length is not important provided that it is not so short that the taper angle becomes excessive: 1½ in (3·72 cm) easily meets this requirement.

In the low-pass filter let $\theta_{1C} = 40°$ and $\theta_{2C} = 20°$; these represent values that are commonly used. Then from eqn. (5) $Z_{01}/Z_{02} = 15{\cdot}6$. The guide wavelength at the low-pass cut-off frequency, for $a = 4{\cdot}17$ cm, is 11·9 cm. The physical lengths for θ_{1C} and θ_{2C} are then found from $b = \theta_c\lambda_g/2\pi$ and are 1·32 cm and 0·66 cm, respectively. These values must later be corrected for the discontinuity susceptance. The filter is to be matched at 4 000 Mc/s and Z_{02}, θ_2 terminations are to be employed. Retaining the convention $Z_{01} > Z_{02}$, eqn. (3) must be rewritten

$$Z_{i2} = Z_{02}\left[\left(\frac{Z_{01} - Z_{02}\tan\tfrac{1}{2}\theta_1\tan\tfrac{1}{2}\theta_2}{Z_{01}\tan\tfrac{1}{2}\theta_2 + Z_{02}\tan\tfrac{1}{2}\theta_1}\right)\left(\frac{Z_{01}\tan\tfrac{1}{2}\theta_1 + Z_{02}\tan\tfrac{1}{2}\theta_2}{Z_{02} - Z_{01}\tan\tfrac{1}{2}\theta_1\tan\tfrac{1}{2}\theta_2}\right)\right]^{1/2} \quad . \quad (3a)$$

The electrical lengths θ_1 and θ_2 at 4 000 Mc/s can be found from the uncorrected lengths and the guide wavelength at this frequency. Substituting for $\lambda_g(=17{\cdot}2\,\text{cm})$ in eqn. (3a),

$$Z_{i2} = 7{\cdot}34Z_{02}$$

Then, from eqn. (9), $b'' = 0{\cdot}79/7{\cdot}34 = 0{\cdot}108$ cm and $b' = 0{\cdot}108 \times 15{\cdot}6 = 1{\cdot}69$ cm.

The correspondence with the guide dimension, b, is fortuitous. The final steps in the design are to correct the lengths for the discontinuity susceptance, and to calculate the length of the high-pass filter and the number of sections in the low-pass filter needed to give the desired attenuations. This method is treated in the original paper.[1]

(4) FILTER CONSTRUCTION

The construction of the filter is shown in Fig. 2. In order to check the design accuracy the model was milled out of solid copper, the filter being made in two sections and then bolted together. The tapers were separately made from the solid and then added to the main assembly. This type of construction is expensive; however, satisfactory filters have been made in wave-

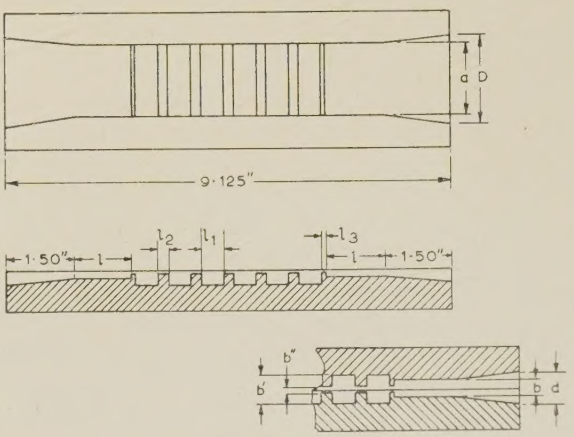


Fig. 2.—Filter construction.

l = 1 in.	D = 2 in.
l_1 = 0·52 in.	a = 1·642 in.
l_2 = 0·194 in.	b = 0·311 in.
l_3 = 0·074 in.	d = 0·666 in.
b' = 0·666 in.	
b'' = 0·042 in.	

guide. For instance, the above design could be made in 2 in × ⅔ in guide quite simply by reducing the broad dimension with inserts along the length, the low-pass sections being formed by inserts across the width. Some simplification, in this case, is effected because b' is the same as the guide dimension b. If desired, this can usually be arranged, but, in general, the practice should be followed with caution because large values of b' lead to errors in the discontinuity-susceptance correction. Designs that lead to large values of discontinuity correction should normally be avoided if experimental modification is undesirable. On the assumption that this correction is moderately small its effect on the image impedance is neglected in eqn. (3). Large values of Z_{01}/Z_{02} are likely to be the chief cause of large discontinuity corrections, and these result from unsuitable choice of θ_{1C} and θ_{2C}. Although governed by spurious pass-band considerations, some freedom of choice should normally be available. Where the latter are particularly rigorous the usual practice of cascading sections with different high-pass cut-off frequencies should be followed.

(5) PERFORMANCE

The performance of the filter is shown in Fig. 3. The agreement with theory is considered good. Filters of this type are basically similar to constant-k low-pass and high-pass networks of lumped-circuit design, and the image impedance therefore varies slowly throughout most of the band, but rapidly near cut-off. Thus, amplitude ripples, especially near cut-off, can be expected. However, by locating the cut-off frequencies somewhat outside the desired pass-band the latter can be kept free from serious ripples. In this example the amplitude of ripples did not exceed 0·25 dB in the chosen pass-band (3 800–4 200 Mc/s); the voltage standing-wave ratio was less than 1·6. The only significant ripple occurs at approximately 4 300 Mc/s and is most likely a result of the impedance mismatch between filter and load just discussed. However, internal reflections in the filter as a result of manufacturing imperfections, or minor errors in the design of the end sections, can also produce noticeable ripples

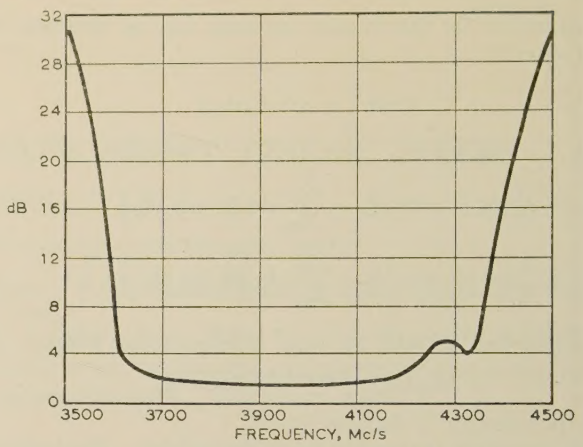


Fig. 3.—Transmission loss characteristic.

(6) ACKNOWLEDGMENTS

The author would like to thank Mr. L. Lewin, for his interest in the work, Mr. H. J. Casford, who constructed the filters, and Standard Telecommunication Laboratories Limited for permission and facilities to publish the paper.

(7) REFERENCES

(1) 'Very High Frequency Techniques', Radio Research Laboratory, Harvard University (McGraw-Hill, 1947), Volume 2, pages 648–740.
(2) Reference 1, pages 689–691, 734.
(3) LEWIN, L.: 'Reflection Cancellation in Waveguides', *Wireless Engineer*, 1949, **26**, pp. 258–264.
(4) Reference 1, p. 674.

PROCEEDINGS OF THE INSTITUTION OF ELECTRICAL ENGINEERS

Part B. RADIO AND ELECTRONIC ENGINEERING (INCLUDING COMMUNICATION ENGINEERING), MARCH 1958

CONTENTS

	PAGE
Radio Observations on the Russian Satellites (Discussion Meeting)	81
The Willis Jackson Report (Summary)	116
Group Chairmen's Addresses	118
Discussion on 'The Control of Nuclear Reactors'	119
The Design of the Control Unit of an Electronic Digital Computer. M. V. WILKES, M.A., Ph.D., F.R.S., W. RENWICK, M.A., B.Sc., and D. J. WHEELER, Ph.D.	121
A Decimal Adder using a Stored Addition Table — M. A. MACLEAN, M.Sc., and D. ASPINALL, B.Sc.	129
An Accurate Electroluminescent Graphical-Output Unit for a Digital Computer. T. KILBURN, M.A., Ph.D., D.Sc., G. R. HOFFMAN, Ph.D., B.Sc., and R. E. HAYES, M.Sc.	136
Discussion on the above three papers	144
Broad-Band Slot-Coupled Microstrip Directional Couplers — J. M. C. DUKES, M.A.	147
The Application of Printed-Circuit Techniques to the Design of Microwave Components — J. M. C. DUKES, M.A.	155
Re-Entrant Transmission-Line Filter using Printed Conductors — J. M. C. DUKES, M.A.	173
Discussion on the above three papers	180
A Coder for Halving the Bandwidth of Signals — A. R. BILLINGS, B.Sc., Ph.D.	182
An Approach to the Design of Constant-Resistance Amplitude Equalizer Networks — J. S. BELL	185
Space Charge as a Source of Flicker Effect — C. S. BULL, Ph.D.	190
An Instrument for the Measurement of Surface Impedance at Microwave Frequencies — A. E. KARBOWIAK, Ph.D., B.Sc.(Eng.)	195
Measurement of Ferrite Loss-Factors at 10 Gc/s — C. M. SRIVASTAVA, M.Sc., Ph.D., and J. ROBERTS, B.Sc.(Eng.)	204
Wide-Band Waveguide Filters with Short Linear Tapers — G. CRAVEN	210

Bibliographical References.—It is requested that bibliographical reference to an Institution paper should always include the serial number of the paper and the month and year of publication, which will be found at the top right-hand corner of the first page of the paper. This information should precede the reference to the Volume and Part.

Example.—SMITH, J.: 'Reflections from the Ionosphere', *Proceedings I.E.E.*, Paper No. 3001 R, December, 1954 (102 B, p. 1234).

Published by The Institution, Savoy Place, London, W.C.2. *Telephone:* Temple Bar 7676. *Telegrams:* 'Voltampere, Phone, London.'
Printed by Unwin Brothers Limited, Woking and London.